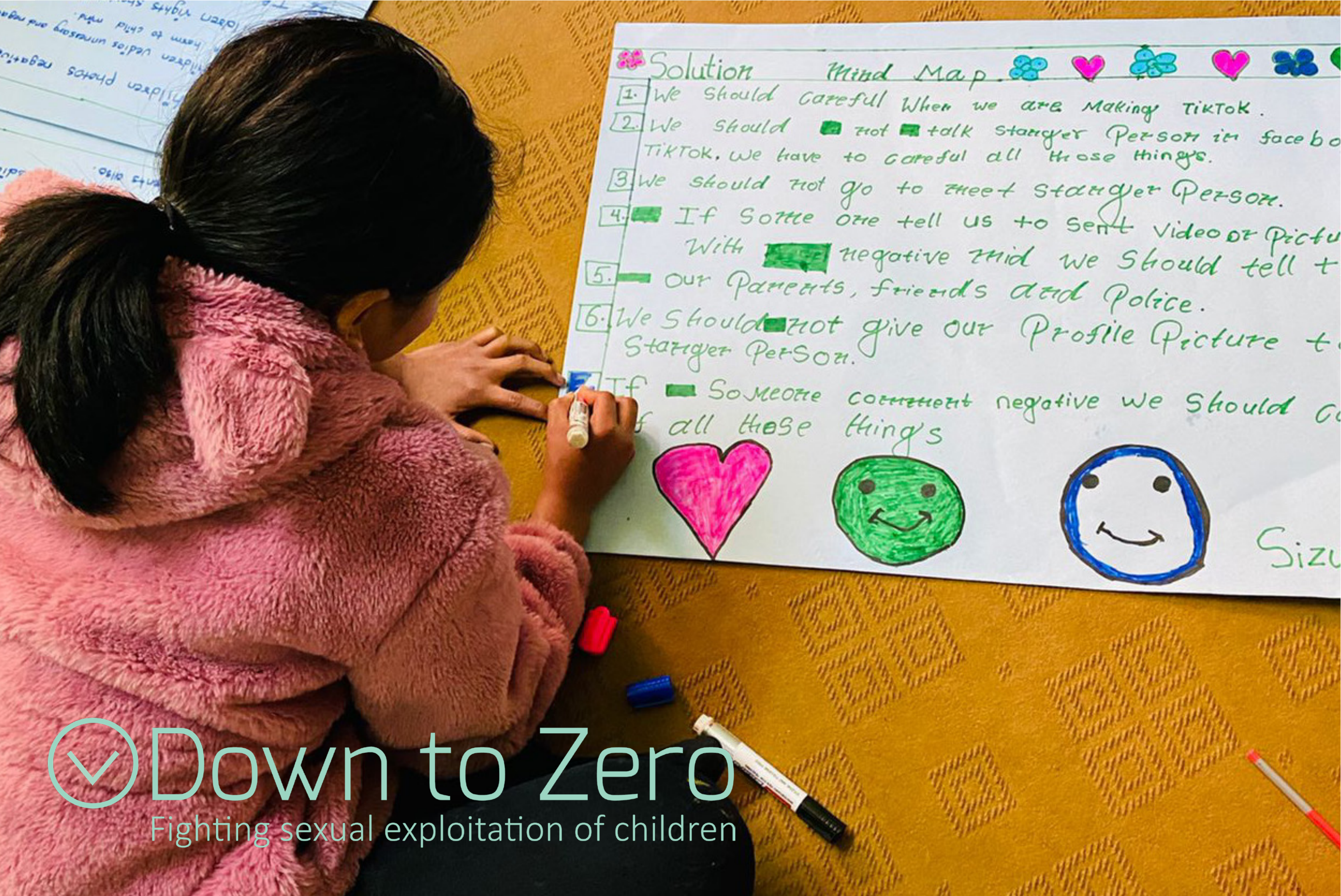
Child safety by design that works against online sexual exploitation of children

Research paper

DOWN TO ZERO ALLIANCE

Prepared by: Nathalie Meurens, Eva Notté, Agata Wanat, Louise Mariano
Reviewed by: Jean Elphick and Hilde Neels
Senior technical experts: Pamela Wisniewski and Jun Zhao
Editor: Siobhan McGonigle



Solution Mind Map

1. We should careful when we are making TikTok.
2. We should not talk stanger person in facebo
TikTok, we have to careful all those things.
3. We should not go to meet stanger person.
4. If some one tell us to sent video or pictu
With negative mid we should tell t
5. our Parents, friends and Police.
6. We should not give our Profile Picture t
Stanger PerSon.
7. If SoMeone comment negative we should c
of all these things

Sizu

Down to Zero
Fighting sexual exploitation of children

This report is part of the Down to Zero Alliance's Building Back Better programme, 2022.

Table of Content

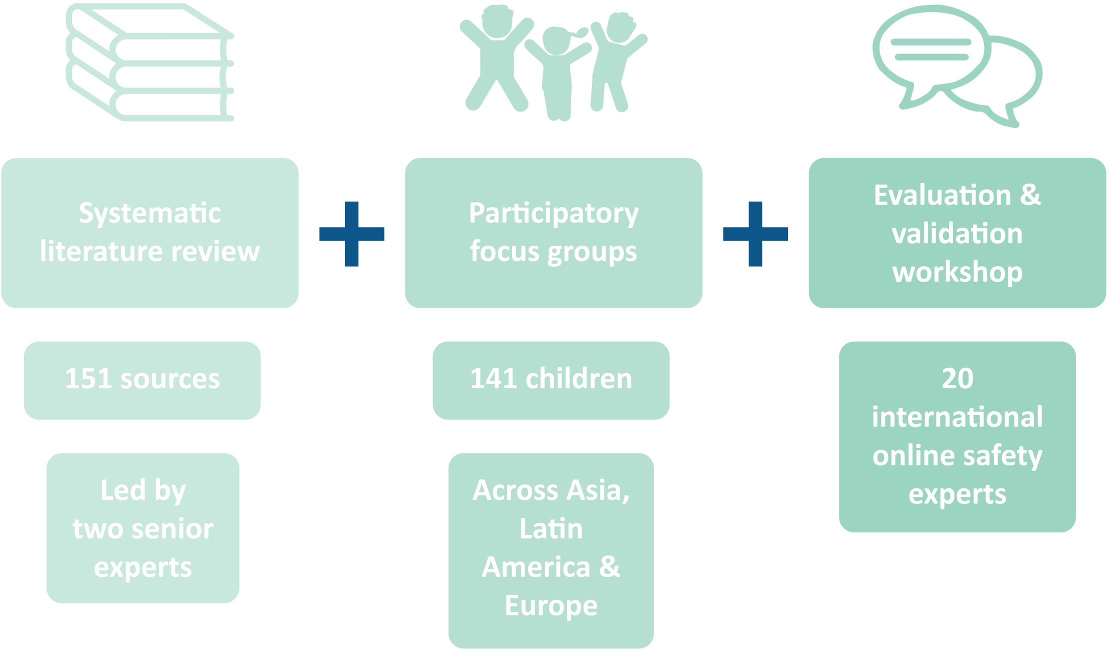
Abbreviations

AI	Artificial Intelligence
AVMSD	Audio-visual Media Service Directive
CJEU	Court of Justice of the European Union
CRC	(United Nations) Committee on the Rights of the Child
CRIA	Child Rights Impact Assessment
CSAM	Child sexual abuse material
CSE	Child sexual exploitation
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
EU	European Union
FGDs	Focus group discussions
GDPR	General Data Protection Regulation
IWF	Internet Watch Foundation
LGBTQI+	Lesbian, gay, bisexual, transgender, queer and intersex
MGRS	Masculine Gender Role Stress
NSPCC	National Society for the Prevention of Cruelty to Children, a UK non-profit organisation
NCMEC	National Center for Missing & Exploited Children, a US non-profit organisation
OECD	Organisation for Economic Co-operation and Development
OSEC	Online sexual exploitation of children
PTSD	Post-traumatic stress disorder
TEU	Treaty on European Union
TFEU	Treaty on the Functioning of the European Union
UNCRC	United Nations Convention on the Rights of the Child
UNICEF	United Nations Children's Fund

Executive summary

The Internet provides a world of opportunity for children and was a lifeline for many during the COVID-19 pandemic, but it also poses a growing risk of exposing children to irreversible harm. Prevalence of Online Sexual Exploitation of Children (OSEC) exploded in 2021, with unprecedented increases in reported cases of grooming, child sexual abuse materials (CSAM), sextortion, and other abuse and exploitation. This study identifies the **most effective child safety by design measures** that can be taken to prevent harm to children online. Protecting children while still enabling their access to, and participation in the online environment is a challenge. In this report, the Down to Zero Alliance's Building Back Better programme provides a clear path through the complexity. Informed by a systematic review of literature, an international panel of 20 senior online safety experts, and focus group discussions with 141 children (aged 11 to 16) in ten countries, a set of concrete safety by design solutions are posed to those with the power to bring about change:

- Five policy recommendations for the **EU** to incorporate child safety by design requirements for social media platforms in upcoming legislation.
- Five solutions for safety by design measures that **industry** should adopt to better serve child users, of which three were also suggested by **children** themselves.
- Four additional solutions developed by **children** themselves, aimed at the **industry**.



Systematic literature review + Participatory focus groups + Evaluation & validation workshop

151 sources

Led by two senior experts

141 children

Across Asia, Latin America & Europe

20 international online safety experts

The focus groups with children took place in Bangladesh, Bolivia, Colombia, Estonia, Nepal, Netherlands, Nicaragua, Philippines, Romania, Thailand.

THE RISK OF CHILDREN BEING EXPLOITED ONLINE IS ESCALATING

Children have the right to benefit from the opportunities of the online world in a safe and non-discriminatory manner. Platforms designed for, or accessed by children should ensure a safe environment. However, research and information published by organisations who investigate and prosecute OSEC indicate that the problem is on the rise. This is due to intersecting factors that exacerbate the risk of OSEC as well as the ripple effect of COVID-19 lockdowns and school closures.

Content risk Children are exposed to or are engaging with potentially harmful violent or sexual content that is not age-appropriate.	Since the start of the COVID-19 pandemic, more children are online, screen time has increased, and children start using the internet and social media at younger ages. For example, studies among children under the age of 13 report that they use social media and messaging apps that are recommended for older children and adults. This means that they can be exposed to content that is not safe or age-appropriate. A recent multi-country study showed that between 54% and 89% of children aged 9 to 11 have seen sexual content online in the past year.	*"It is right to lie about age online because young people can't access Facebook, YouTube, TikTok and online gaming"* (Focus group discussions child from Nepal)
Contact risk Children are experiencing or are being targeted by potentially harmful contact online.	Being 'befriended' or groomed online leads to new risks of harm when communication between children and potential offenders is established on apps that children enjoy and spend time on. To illustrate, apps popular among children including Facebook-owned applications and Snapchat are the platforms used in over 70% of online grooming cases. This means that children are exposed to risky contact during critical periods of their development, often before their cognitive, risk-detection, self-regulation, and coping skills have matured.	*"There are adults that talk to teenage girls, they bother them, tell them they have work for them, if they want work, and different things I have seen."* (focus group discussions child from Bolivia)
Conduct risk Children witness, participate in or are victims of potentially harmful conduct.	Children engage in risky behaviour as they underestimate the risks associated with the creation or sharing of sexually explicit messages and content. A survey of law enforcement officials in 39 countries reported a significant increase in OSEC since 2020. Multi-country research showed an increased prevalence of online grooming, sexual extortion, live-streaming, and self-generated CSAM even before the COVID-19 pandemic. In 2021, reports of CSAM escalated radically, most commonly featuring girls and young children under the age of 13.	*"When we talk with [unfamiliar people online], we would talk in a way as we usually and normally do with others- regardless of how much we know of them. For this reason, we may innocently become victims because we cannot keep up with a trick the online friend may put on us."* (focus group discussions child from Thailand

This means that OSEC is an escalating problem and that regulators and the tech industry are failing to take responsibility for users' safety and wellbeing. Social media and other online spaces are not adequately designed to keep children safe, putting girls and young children particularly at risk.

Children are a diverse group and intersecting factors contribute to risks, in addition to age and stage of development. OSEC is a form of **gender-based violence.** Hegemonic gender norms dictating that boys and men be physically and emotionally tough and heterosexually dominant over girls and women result in adolescent girls being most vulnerable. Girls are encouraged to be submissive, whereas boys are under pressure to perform assertive masculine roles. Gender therefore foreshadows online behaviour and predicts both the risk of experiencing victimisation and the risk of perpetrating violence. Beyond gender, factors such as disability, sexuality, race and ethnicity, mental health, and family dynamics among others influence risk of sexual harm online.

PLATFORMS ARE NOT TRANSPARENT ABOUT THEIR SAFETY MEASURES AND THEIR EFFECTIVENESS

Some safety measures are in place on various platforms popular among children, but it is difficult to know to what extent or how they operate. As there is a lack of transparency of the complete suite of measures in use, the table below represents publicly available data only.

	FACEBOOK & INSTAGRAM	SNAPCHAT	TIKTOK	YOUTUBE
Minimum age	13	13	13	13
Age verification by self-declaration	✓	✓	✓	✓
Age verification using various sources upon registration	✗	✗	✗	✗
Age assurance methods after registration	✓	✗	✓	?
Separate platform for children	✗	✗	✗	✓
Default privacy default settings	✓	✓	✓	✗
Extra default privacy settings for children	✓	✗	✓	✓
Use of classifiers	✓	?	✓	✓
In-app reporting	✓	✓	✓	✓
Referral of reports to other organisations	✓	?	?	✓
Deterrence and warning messaging	✓	?	?	?

As a legal principle, any restrictions to the child's freedom of expression in the digital environment, such as filters and other safety measures, should be established in law, should be as assessed as necessary and proportionate, and should be transparent and well communicated to children in age-appropriate language. This is not the case at present. For the most part, we are either in the dark about what safety features are being used, or aware of the ineffectiveness of their patchy implementation by some platforms. The children in the focus groups confirmed that they encounter a lot of content they feel should not be seen by children, including violence and pornography. In another focus group exercise, a grooming chat was displayed. Children had first-hand experience of similar chat conversations or had heard about these from friends.

"There are so many things on the internet that children shouldn't watch."
(Focus group discussions, child from Nepal)

"There is a lot of content on the internet that can have a negative impact on us."
(Focus group discussions, child from Romania)

Children corroborated evidence shows that **current measures are not sufficient to shield children** from encountering potentially harmful situations in their social media use. Furthermore, rather than designing preventive features, platforms are reactive, often only taking remedial action once harm is done. Companies tend to prioritise profit over online safety by using technologies that not only fail to prevent harm, but that actively facilitate risk exposure.

Designing platforms to attract users, generate engagement, and encourage interactions can be in tension with promoting the safety and wellbeing of users. Very few industry leaders are transparent about how their algorithms shape user experiences, however, the following pose threats to child safety online:

- **Cognitive biases and design tactics** (also called **dark patterns**) are used to nudge and manipulate users to make risky choices and to share more information. For example, some platforms make it more difficult for users to choose stronger privacy settings by providing an overwhelming over-choice of privacy options. Other nudging techniques fuel contact risks by recommending adult friends or followers to children or by including location tracking in some 'friend suggestion' systems.
- **Third-party libraries** collect sensitive data about online users, including children, which is then used for data profiling and targeted advertising. Children may agree to terms that allow personal data, call logs, browser history, and contact information to be shared.
- **Terms of Use** agreements and Codes of Conduct are usually long, mainly textual, and not child friendly. The text itself is technical, making it not very accessible to children. As a result, children do not know what they have agreed to.
- There is an **absence of guidelines for developers** to ensure child safety.

CHILDREN ARE BOTH PART OF THE PROBLEM AND PART OF THE SOLUTION

Children, especially teenagers, are risk-takers and overestimate their ability to cope with risks. Even when they understand the risk, teenagers in particular will prioritise the social benefits over their safety. Teenagers explore their sexuality online through sharing, and over-sharing, private and at times intimate content. Sexting can be grounds for harmful behaviour, especially when one party is pressuring the other for content and/or overlooking the need for consent.

Children themselves may commit OSEC against other children, including through self-generating CSAM, grooming or sexual harassment. Research indicates that a substantial percentage of child sexual harm and abuse is committed by people under the age of 18, including online.

Yet children are also part of the solution as they can be involved in identifying risks online and can learn how to deal with the risks they face. Effective approaches to keeping children safe must recognise that **children exercise agency** online. They can take risks, yet they are also capable of empowered online behaviour as they grow and mature.

Research shows that over-restrictive parental controls or surveillance and fear-based digital abstinence approaches are not as effective as technological design solutions that guide children in dealing with the risks they face online. To support the development of their skills and supplement education and supervision, **child safety by design** is a recommended strategy that incorporates features that both **empower** children and safeguard user safety and wellbeing through the design of apps and digital platforms.

> *"Sometimes there are advertisements of adult videos or content, but we don't watch it because I know it is not safe for children. I knew this from family, friends, and teachers. There are no online safety classes in school. We get little information on social and computer classes."*(Focus group discussions, child from Nepal)

CHILD SAFETY BY DESIGN OFFERS SOLUTIONS THAT COUNTER OSEC

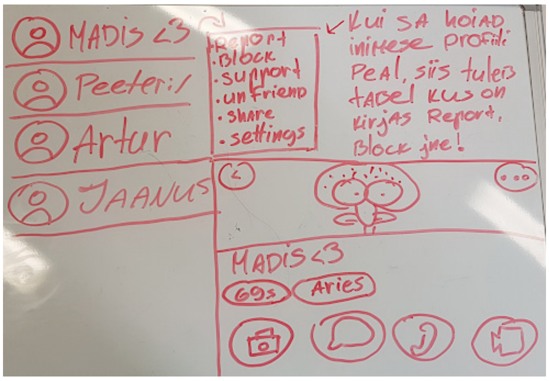
Safety by design is a user-centred approach that puts user safety and rights at the core of the design and development of services and products. These design features can help prevent OSEC by excluding predators from children's online forums and ensuring age-appropriate online experiences for young users.

While some online platforms and social media apps have started addressing child safety, they fall short in implementing a holistic set of effective measures to prevent the exploitation of users.

Ensuring safety online for children requires striking a **balance between their protection and the full realisation of their right to freedom of expression and right to privacy**. The child's right to freedom of expression and privacy are embedded in various children's rights conventions (such as in Article 13 and 16 of the United Nations Convention on the Rights of the Child (UNCRC)). As children continue to seek independence in the digital world, there is often a blur when it comes to the role of adults interfering with the child's online presence.

Children in the focus groups supported platforms offering **safety features that align with their age**. Because younger children have different needs to older children, platforms should deploy parental control features that are age-appropriate. In line with this, parental monitoring and restriction should be limited to children under the age of 13. For parents of teenagers over 13, **privacy-preserving parental control** apps should only provide high-level category information rather than details. Age verification and assurance can be strengthened by using officially provided, automatically generated, user-reported and third-party data but privacy needs to be kept in mind. Safety features for **adolescents** (13 to 17 years old) should empower them to protect themselves by encouraging **self-regulation**.

Children and the expert panel supported literature that recommends that **intelligent privacy by default** should become the industry standard. These default settings for children's accounts should include friends-only permissions and should disable geo-tagging. In addition, retroactive privacy features, such as the ability to untag, delete, block, and report inappropriate content promotes shared responsibility and agency.



Source: Focus group in Estonia

Children suggested some more immediate ways of strengthening the design of platforms. These included better **visibility** of rules and making it easier to report violations by providing many clearly visible options. Unsurprisingly, children raised the issue of imposing consequences on abusers. They also supported the fact that platforms should play a bigger role in keeping children safe by using evidence-based computational risk detection and mitigation strategies, like intervening with pop-up messages to prompt children to be aware of and respond to risks.

In addition, **knowledge is a powerful preventive tool**. Knowledge on the patterns and risks of OSEC could in itself also contribute to more caution for multiple stakeholders:

- For designers of online platforms so they know where their focus of safety design should lie.
- For children so they can protect themselves better, recognise red flags, and know what inappropriate behaviour is.
- For adults (parents and professionals) so they know what to teach their children and which signals they could spot when children face online harm.
- For policy-makers so they adopt laws that can effectively protect children online.
- For researchers so they can identify areas for further research and provide evidence-based effective solutions.

WE NEED TO TAKE THE OPPORTUNITY TO INCLUDE CHILD SAFETY BY DESIGN AS A LEGAL REQUIREMENT

The EU policy framework already establishes some minimum standards for fighting OSEC, yet EU legislation is largely insufficient to tackle the scale and complexity of the issue. It does not incentivise social media platforms to put in place effective safety measures and the legal provisions are too vague to lead to effective safety measures. **Without regulation, online platforms will continue to compete for profit at the expense of user safety**.

The EU framework has incorporated some elements of safety by design in its requirements, including transparent and user-friendly reporting mechanisms, age verification, and parental control systems. The Audio-visual Media Service Directive (AVMSD) is a key instrument to that end. Yet, it is limited in scope and lacks effective requirements.

The EU does have the power to enact stronger policies and, in representing 27 Member States, is well positioned to adopt **global standards** that can protect children worldwide. Several pieces of EU legislation and regulations are due for revision or are in proposal stage, representing an opportunity to incorporate recommended safety by design features.

EVIDENCE-BASED AND CHILDREN-INSPIRED RECOMMENDATIONS AND SOLUTIONS

EU RECOMMENDATION 1

- **Decriminalise consensual sexting** among minors in its revision of the Child Sexual Abuse Directive

Puberty is a stage of sexual exploration and development. Online platforms can provide a private outlet for such exploration. The problem is that teenagers are often unaware of, or underestimate the risks of their online behaviours, like sexting. The line between consensual (legal) sexting and risk behaviours or even OSEC is blurry. Currently, the legislation does not reflect the complexity of the phenomenon of sexting. There is, thus, a **gap between law and practice**.

Protecting children from OSEC should be a learning, supporting, open, and continuous process between children, parents, legislators, and online service providers. The EU must ensure that children can safely explore their sexuality without fear of a criminal response, yet harmful illegal behaviour must be addressed.

EU RECOMMENDATION 2

- Require platforms to have **effective age verification systems**

INDUSTRY SOLUTION 1

- Strengthen **age verification and assurance** by using multiple sources
- Children also recommended to use **official documents or parents** to verify age and identity

Most social media applications resort to self-declaration of age which, while cheap and easily implemented, cannot be considered an effective age assurance mechanism. Children under the age of 13 are likely to lie about their age, knowing that otherwise they would be excluded from accessing the service. This leads to many users getting access to content that is not meant for them yet.

For providing children with an age-appropriate experience, **age verification and assurance** should be strengthened by using multiple sources. This could be officially provided, automatically generated, user-reported and third-party data.

Children expressed that anonymous and fake accounts should no longer be allowed. Many children in the focus groups therefore indicated that age and identity of all potential users should be checked before users are allowed on platforms. The children advised that a users' identity card or passport should be uploaded and checked. The identity of parents could be checked as part of the registration process for children.

EU RECOMMENDATION 3

- Require all platforms to have **transparent reporting mechanisms and referral systems** for children reporting OSEC

CHILDREN'S SOLUTION 1

- Increase the visibility of rules and reporting mechanisms

CHILDREN'S SOLUTION 2

- Introduce stronger reporting mechanisms and consequences for abusers

Terms of Use agreements and Codes of Conduct are usually not child-friendly. Long documents of fine print with technical jargon are not accessible to most users, let alone children. The default action is often to click 'Agree' without knowing what is being agreed to. When things go wrong, and something uncomfortable happens online, children do not always share or report the incident, often indicating that they are **unaware of how reporting works**.

The EU should make it more clear what transparency means and should require platforms to **enhance transparency** of reporting mechanisms. Platform holders should be required to support the reporting process for their users. There should always be a clearly visible reporting button, where various options for reporting are presented. Platforms should also be required to have appropriate aftercare, for instance with a time limit for updating the person who reported about the process or outcome or with a referral to other organisations that can provide appropriate services.

Many children mentioned that visibility of house rules, but also of reporting options should be increased. They suggested having the rules visible at all times in a **sidebar**. When they would come across something distressing, they mentioned that reporting or danger buttons should be visible on screen. This would lower the threshold to report incidents. The children also found that when they reported, nothing ever changed. They therefore suggested stricter consequences for people that violated the rules, such as a public warning on profiles or removing the user after multiple strikes.

INDUSTRY SOLUTION 2

- Develop features that encourage **child/teen empowerment and self-regulation**

INDUSTRY SOLUTION 3

- Deploy parental control features adapted for children across all age groups

Overly restrictive parental control and fear-based abstinence approaches may **reassure parents, but they are not effective**. Limiting access to the digital world and opportunities erodes learning, relationship-building and self-expression. Parental control apps can limit screen time and mitigate some risks of being exposed to harmful content, however, either children are tech savvy enough to get around them or the controls are overly restrictive. Such approaches can also damage trust between parents and children. Instead, children report needing support from trusted adults to deal with online risks.

Parental monitoring and restriction should be limited to children aged 3 to 12 years old, while category-based, privacy-preserving parental control apps are a better fit for adolescents aged 13 to 17 years old when they recognise their needs for autonomy and agency. For adolescents, safety features should **empower children to protect themselves** when using technologies through self-regulation and self-monitoring, promoting shared responsibility, trust, and communication between parents and children. By taking a more 'teen-centric' rather than 'parent-centric' approach to adolescent online safety, designers can help teens foster a stronger sense of personal agency for self-regulating their own online behaviours and managing online risks.

INDUSTRY SOLUTION 4

- Develop features that encourage **child/teen empowerment and self-regulation**
- **Children also recommended differentiating the design for children**

Cognitive biases and **design tactics** (dark patterns) are often used by platforms to manipulate users into sharing more information, stay active on the platform for longer, and encourage reactions and interactions. Techniques include nudging, leading language, instant gratification linked to sharing data, or paradoxically providing an overwhelming over-choice of privacy options.

Platforms should deploy **intelligent privacy features** to ensure the highest privacy settings for any child and make settings customisable for certain age groups. Those features should include intelligent privacy settings by default for all user accounts under 18 years of age, with proactive privacy features as well as reactive privacy features. These should not be a luxury. All businesses should be required to implement regulatory frameworks, industry codes, and terms of services that adhere to the highest standards of ethics, privacy, and safety in relation to the design, engineering, development, operation, distribution and marketing of their product and services.

In the focus groups, children also acknowledged the importance of differentiating between adults and children. Some children suggested having different designs for younger and older children as well.

EU RECOMMENDATION 4

- Make online platforms legally accountable for **establishing minimum safety standards** to keep children safe

CHILDREN'S SOLUTION 3

- Take extra measures for content involving children

Designers face a number of **challenges** such as the **lack of awareness and/or guidelines** on ensuring child safety. The Audio-visual Media Service Directive (AVMSD) requires video-sharing platforms to ensure that children may not 'normally hear or see' content that may impair their physical, mental, or moral development. It does not, however, define what type of content falls in this category and these requirements are largely insufficient to tackle OSEC.

Legislation should be put in place to make online platforms accountable for establishing **minimum safety standards** to keep children safe from OSEC and CSAM that occur on their platforms, including carrying out regular Child Rights Impact Assessments (CRIAs).

Children suggested that **extra precautions** should be taken to make sure that their information and content is not misused or used without their consent. This could be done through disabling downloads or screenshots of information about children.

INDUSTRY SOLUTION 5

- Implement technology that can **detect risks** combined with **risk mitigation** strategies
- Children also recommend platforms to have a bigger role in keeping them safe by having moderators and intervention messaging

High-quality computational **risk detection tools** are currently under development, but they still face many challenges and technical issues. Most of the tools are developed at software level only, meaning they can only apply to images and videos that have already been created and disseminated. There is a lack of focus on using such tools at hardware level (i.e. device level) to act as a preventive measure in the case of sexting or OSEC.

Evidence-based computational **risk detection** combined with **risk mitigation** strategies should become the industry standard in order to help identify risks and prompt children to respond to those risks. Computational risk-detection tools should be grounded in evidence and use a child-centred design approach.

Education and awareness warnings and prompts should be used by service providers to offer in-the-moment information to parents and children on how to respond in the context of a risky situation or experience.

Children indicated that more can be done by platforms to make sure that there is age-appropriate content and conversations. They therefore suggested moderates that monitor the platform. The children also thought it was helpful to have **pop-ups** and warnings when accessing specific content or when sharing content. This would make them rethink certain actions. Another idea offered was a ‘think before you click’ advertisement.

EU RECOMMENDATION 5

- Support and strengthen initiatives aimed at **education, awareness, action research**, and **offender interventions**

CHILDREN’S SOLUTION 4

- Use **popular media and well-known people** to deliver safety messages

Education programmes about sexuality as well as sexual and reproductive health rights tend not to be sufficiently comprehensive or are simply lacking in terms of consent, exploring sexuality online, and gender sensitivity. The **stigma** of such topics often means that adolescents learn from their peers, media, online platforms, or pornography, with high risks of perpetuating sexist and violent behaviours that they have internalised. The **fear of stranger danger** is overestimated, as young people are more often victimised by people they know, including by children themselves.

Education is an important measure for preventing and tackling OSEC. Instead of shielding children from risks altogether, it recognises children’s need to be online and teaches them how to recognise and deal with risks, thereby building digital **resilience.** Teaching children **digital literacy** should encompass a broad range of knowledge and skills that go beyond just the technical and operational aspects, such as self-awareness online, empathy, and communication. Online safety training should be incorporated into education about **real-world risks**, as similar skills and knowledge are needed to prevent and tackle child abuse both offline and online. **Professionals and parents** should be equipped with sufficient tools and knowledge to be able to teach children about online safety in an age- and development-appropriate manner.

Children proposed using well-known public figures, such as **influencers**, to deliver safety messaging to make it appealing to children. Messaging could be delivered through the platforms themselves, via news programmes, and on television.

In **conclusion**, safety by design is complex and complicated, as there are many associating and intersecting factors that need to be incorporated for design to be effective. Online platforms and their design could play a significant role in keeping children safe online, but the measures currently in place are not sufficient. EU policies are also not stringent enough to force online platforms to change their designs. To contribute to better protecting children against OSEC, this research report provides a deeper understanding of the problem and gives concrete solutions linked to EU and industry policy recommendations. This report highlights the effectiveness and potential of the safety by design approach not only to stop OSEC, but also to strengthen children’s digital resilience, valuing their autonomy and freedom of expression and privacy. The report also aims to tackle the problem in a concrete and sustainable way by supporting the revision of the EU legal framework, so that children can benefit from what the internet offers in a safe and age-appropriate manner.

1. Introduction

KEY MESSAGES

- More and younger children are **increasingly active online** in the wake of COVID-19 restrictions and school closures
- Prevalence of online sexual exploitation of children (OSEC) **exploded in 2021**, with unprecedented increases in reported cases of grooming, child sexual abuse material (CSAM), sextortion, and other abuse and exploitation
- **OSEC is harmful** to children's dignity, development, health, and survival, and the nature of the internet makes it difficult to delete CSAM and prevent retraumatisation
- Social media and other online spaces are **not adequately designed** to keep children safe, putting girls and young children especially at risk
- Existing online safety measures employed by digital platforms and social media are **ineffective and reactive**, often only taking remedial action once harm is done
- The present study aims to further understand child safety by design and how it can help better protect children online against OSE
- This study is based on an extensive **literature review** of 151 sources, mostly peer-reviewed articles and key reports, **focus groups** with children in 10 countries across the world, as well as **input from experts**

As more and more children become active online and begin using social media at younger ages, the responsibility to keep them safe online grows. In the aftermath of the COVID-19 pandemic, **68% of law enforcement officials surveyed in 39 countries reported an increase in Online Sexual Exploitation of Children (OSEC)**.[1]

OSEC includes *"all acts of a sexually exploitative nature carried out against a child that have, at some stage, a connection to the online environment. It includes any use of ICT that results in sexual exploitation or causes a child to be sexually exploited or that results in or causes images or other material documenting such sexual exploitation to be produced, bought, sold, possessed, distributed, or transmitted"*.[2] This research supports a call for the European Union (EU) to adopt stronger child safety by design requirements, to prevent harm caused by OSEC.

Child safety by design can be understood as an approach to tackling online risks that focuses on anticipating harms and ensuring that preventative measures, or steps to avoid or minimise risks, are embedded in the design, development and deployment of online and digital services and products.

By combining in-depth desk research with safety by design solutions suggested by children from 10 countries and a panel of international experts in online safety, the findings are presented as evidence-based policy recommendations to:

- ensure children's online access in a safe way;
- effectively prevent OSEC;
- address shortfalls in current online safety approaches, which are proven ineffective;
- promote user privacy, in line with the General Data Protection Regulation (GDPR); and
- respect children's rights by balancing their rights to protection and their rights to participation in the online environment.

[1] NetClean (2021) NetClean Report – COVID-19 Impact 2020 – A Report about child sexual abuse crime. Available at: https://www.netclean.com/wp-content/uploads/2021/01/NetCleanReport_COVID19_Impact2020_pages.pdf (Accessed 18 April 2022).

[2] ECPAT International (2016) Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse (commonly known as the 'Luxembourg Guidelines'), p.17. Available at: https://ecpat.org/wp-content/uploads/2021/05/Terminology-guidelines-396922-EN-1.pdf (Accessed 18 April 2022).

This research is part of the Building Back Better programme of the Down to Zero (DtZ) Alliance.

Down to Zero Building Back Better is a partnership with the Dutch Ministry of Foreign Affairs and DtZ Alliance members: Terre des Hommes (lead), Defence for Children – ECPAT, Free a Girl, ICCO (part of Cordaid), and Plan International NL. It aims to end sexual exploitation of children in 12 countries (Asia and Latin America) by addressing all interrelated actors: children, community, government, law enforcement, and the private sector. The programme ran from May 2021 until May 2022.

1.1 WHAT IS CHILD SAFETY BY DESIGN?

Child safety by design places the **safety of users at the centre** of the design of online services, while placing responsibility for users' safety on online service providers.

> **Child safety by design** consists of *"taking preventative steps to ensure that known and anticipated harms have been evaluated in the design and provision of an online service; that user empowerment and autonomy are secured as part of the in-service experience; and that organisations take ownership and responsibility for users' safety and well-being, and are clear about the steps required to address any issues"*.[3]

In its General comment No. 25, the UN Committee on the Rights of the Child recognises safety by design as a necessary approach to the full protection of children's rights. Accordingly, safety by design should be integrated in the design of digital services and products that children use. The UN CRC specifies that safety and protective measures should take into account children's evolving capacities.[4]

Australia's eSafety Commissioner has defined safety by design principles to include:[5]

- **Service provider responsibility**: Service providers should share the burden of safety with the users. Service providers are responsible for ensuring that known and potential harms are identified (e.g. through risk assessment) and addressed in the design and development of their services.
- **User empowerment and autonomy**: Products and services should align with the best interests of users, whereby features and functionality are designed to protect human rights. Platforms and services need to engage in meaningful consultation with diverse and at-risk groups to ensure their features and functions are accessible to all and protect their rights.
- **Transparency and accountability**: Companies should publish data and information on how they enforce their own policies and on the efficacy of safety features and innovations. Where safety innovations are proven effective in improving user safety and deterring online abuse, they should be shared widely.[6]

Safety by design can be used as an effective method of countering OSEC through techniques such as age assurance, intelligent privacy default, filters, risk detection, parental controls, and robust reporting mechanisms to ensure "age-appropriate online experiences".[7]

[3] UNICEF (2021) Digital Age Assurance Tools and Children's Rights Online across the Globe. Available at: http://c-fam.org/wp-content/uploads/Digital-Age-Assurance-Tools-and-Children-s-Rights-Online-across-the-Globe-1_LT.pdf (Accessed 18 April 2022).

[4] UN CRC (2021) General comment No. 25 (2021) on children's rights in relation to the digital environment. Available at: https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation (Accessed: 27 April 2022)

[5] eSafety Commissioner Australia (2019) Principles and background. Available at: https://www.esafety.gov.au/industry/safety-by-design/principles-and-background (Accessed 18 April 2022).

[6] ibid

[7] WeProtect Global Alliance (2021) Global Threat Assessment 2021 – Working together to end the sexual abuse of children online, p.39. Available at: https://www.weprotect.org/global-threat-assessment-21/#report (Accessed 18 April 2022).

The approach encompasses specific techniques to ensure safety (e.g. age verification, strong privacy settings by default, and hiding children user accounts from engine searches, etc.) as well as an overall approach to design that places children's needs, rights, and safety at the core by assessing their needs and the risks they face, and by ensuring compliance with children's rights at all stages. To be effective, safety by design needs to account for the specific nature of the risks it tries to tackle, including gender and intersectional dimensions as well as the impact of age. This can be done through risk assessments and by including children in the design process, for instance.

Safety by design offers an attractive approach to better protecting and empowering children online. This user-centred approach puts children's safety and rights at the core of the design and development of services and products. **Safety by design features can help exclude predators from children's forums, and ensure age-appropriate online experiences for young users**.[8]

1.2 WHY IS IT NECESSARY TO REDOUBLE EFFORTS TO KEEP CHILDREN SAFE ONLINE?

1.2.1 More children are active online from a younger age

The COVID-19 pandemic redirected social and academic life online for many children. It is estimated that **children constitute one third of internet users** worldwide[9] and over **800 million children are active on social media**.[10] Internet Watch Foundation (2021) suggested that COVID-19-related lockdowns contributed to 'younger and younger children' spending an extensive amount of time online and being targeted by groomers on a large scale.[11]

Reports from all corners of the globe suggest that young children are using social media platforms despite not reaching the platforms' minimum age requirements. Cyber Safe Kids reported that 84% of surveyed 8- to 12-year-olds were using social media and messaging apps in 2020.[12] In the United States, one third of TikTok users were under 14 years old in 2020.[13] According to EU data from 2020, the number of 9- to 11-year-olds who reported visiting a social networking site daily ranged between 11% in Germany and 45% in Serbia.[14]

Being online and on social media provides tremendous opportunities for children. It has become a place for learning, play, and engaging in social interactions, keeping up with friends and families.

"I like to chat with different people, you can make great friends [online]."
(Focus group discussion, 15-year-old girl, Bolivia)

"I can see the photos of everyone on Facebook, especially my relatives who are far away from me."
(Focus group discussion, 14-year-old boy, Nepal)

"I spend time on YouTube because I look for information for the homework I have to do for school."
(Focus group discussion, "Gringo", Nicaragua)

Being active online also means facing risks of harms such as cyberbullying, grooming, and sexual extortion. If risk is calculated by considering the likelihood of a negative occurrence in combination with the impact, having more children spend increasing lengths of time online at a younger age increases risk.

[8] ibid

[9] UNICEF (2019) Growing Up in a Connected World: Understanding Children's Risks and Opportunities in a Digital Age. Available at: https://www.unicef-irc.org/growing-up-connected (Accessed: 18 April 2022).

[10] End Violence Against Children (n.d.), Safe Online. Available at: https://www.end-violence.org/safe-online (Accessed: 18 April 2022).

[11] Internet Watch Foundation (2022) Three-fold increase of abuse imagery of 7-10-year-olds as IWF detects more child sexual abuse material online than ever before. 13 January. Available at: https://www.iwf.org.uk/news-media/news/three-fold-increase-of-abuse-imagery-of-7-10-year-olds-as-iwf-detects-more-child-sexual-abuse-material-online-than-ever-before/ (Accessed 18 April 2022).

[12] Cyber Safe Kids (2021) Annual Report 2020. Available at: https://www.cybersafekids.ie/wp-content/uploads/2021/09/CSK_Annual_Report_2020_web.pdf (Accessed 18 April 2022).

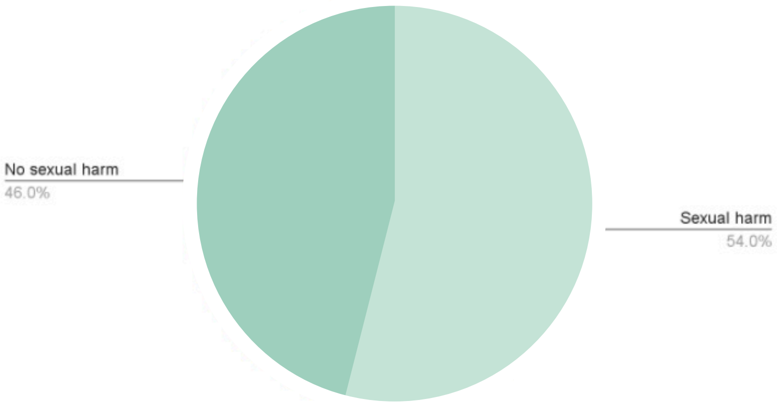
[13] New York Times (2020) A Third of TikTok's U.S. Users May Be 14 or Under, Raising Safety Questions. 14 August. Available at: https://www.nytimes.com/2020/08/14/technology/tiktok-underage-users-ftc.html (Accessed 18 April 2022).

[14] Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S., and Hasebrink, U. (2020) EU Kids Online 2020: Survey results from 19 countries. https://doi.org/10.21953/lse.47fdeqj01ofo.

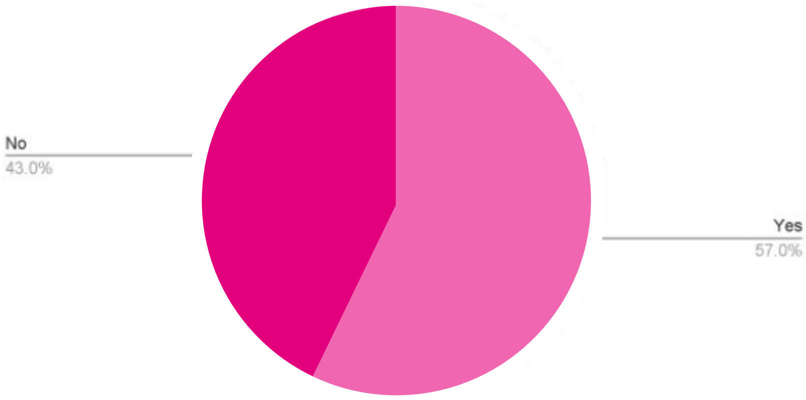
1.2.2 OSEC is on a steep upward trajectory

In addition to the increasing likelihood that children might encounter OSEC or CSAM online, the negative impact of this is the harm caused. In a global survey, **54% of young people (57% of all girls; 48% of all boys) reported having experienced online sexual harms before they were 18 years old**, including within interaction with adults and being asked something sexually explicit or sent sexually explicit content.[15]

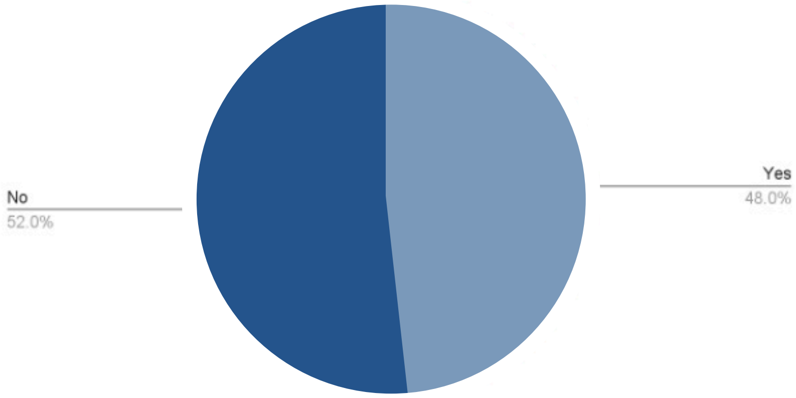
Sexually harmed before 18



Percentage of girls having at least one online sexual harm before 18



Percentage of boys having at least one online sexual harm before 18



[15] Economist Impact and WeProtect Global Alliance (2022) Estimates of childhood exposure to online sexual harms and their risk factors. Available at: https://www.weprotect.org/economist-impact-global-survey/#report (Accessed 18 April 2022).

CSAM refers to images and videos *"that depicts and/or that documents acts that are sexually abusive"*.[16]

The UN CRC's General comment No. 25 (2021) on children's rights in relation to the digital environment categorises four kinds of online risks children are susceptible to: content, contact, conduct and contract risks.[17] The first three types of risks encompass violent and sexual content, exploitation, and abuse (including SEC)[18].

Table 1. Overview of online risk categories

Risks for Children in the Digital Environment				
Risk Categories	**Content Risks**	**Contact Risks**	**Conduct Risks**	**Consumer Risks**
Cross-cutting Risks	Privacy Risks (Interpersonal, Institutional & Commercial)			
	Advanced Technology Risks (e.g. AI, IoT, Predictive Analytics, Biometrics)			
	Risks on Health & Wellbeing			
Risk Manifestations	Hateful content	Hateful Encounters	Hateful Behaviour	Marketing Risks
	Harmful content	Harmful Encounters	Harmful Behaviour	Commercial Profiling Risks
	Illegal content	Illegal Encounters	Illegal Behaviour	Financial Risks
	Disinformation	Other Problematic Encounters	User-generated Problematic Behaviour	Security Risks

Source: Children in the digital environment: Revised typology of risks – OECD Digital Economy Papers[19]

[16] ECPAT International (2016) *Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse*, p.39.

[17] UN CRC (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment*.

[18] Hasebrink, U., Livingstone, S. and Haddon, L. (2008) Comparing children's online opportunities and risks across Europe: cross-national comparisons for EU Kids Online. Available at: http://eprints.lse.ac.uk/21656/1/D3.2_Report-Cross_national_comparisons.pdf (Accessed 27 April 2022).

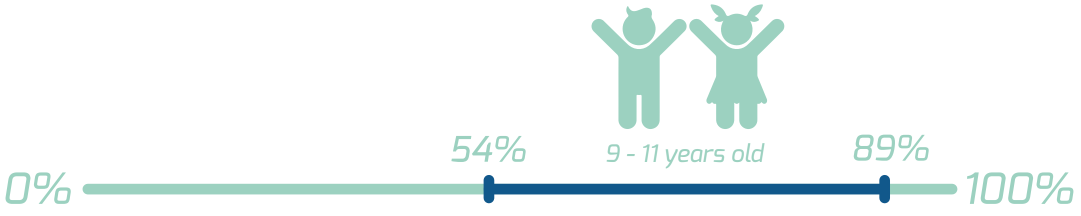
[19] OECD (2021) Children in the digital environment: Revised typology of risks – OECD Digital Economy Papers, No. 302, Figure 1, p.7. https://doi.org/10.1787/9b8f222e-en.

Content risks refer to children being exposed to, or engaging with potentially harmful violent or sexual content. Children may encounter sexual content either by accident, upon recommendation by friends, or receive such content directly from strangers or someone they know.

> *"When I was seven [years old] some man sent me pictures. I showed them to my father. He shared it on FB and asked what to do in this situation. We blocked the man as well."*
> (Focus group discussion, 12-year-old girl, Estonia)
>
> *"There was (were) a stranger(s) approaching me via online video call on Facebook, which I then immediately blocked that person."* (Focus group discussion, Thailand)
>
> *"I was connected with a female friend on Facebook. After some days, I found [out that he] was a boy which made me tremendously embarrassed. I eventually blocked the person."*
> (Focus group discussion, 13-year-old girl, Bangladesh)

Analysing data about 8- to 12-year-olds from 30 countries, Park et al. (2020) showed that children who own a smartphone and reported high social media and gaming activity had an **89% chance of exposure to potentially harmful content** from at least one risk category,[20] such as sexual content. 5Rights Foundation has found that children registering a new social media account are likely to receive inappropriate content within as little as 24 hours after setting up the account.[21]

A multi-country survey carried out by the United Nations Children's Fund (UNICEF, 2019) showed that a high percentage (between 11% and 55%, depending of the country) of children have encountered sexual images online and many faced online harm, especially older children (between 15 and 17 years old)[22]. A meta-analysis of available data suggested that **one in five young people experience unwanted online exposure to sexually explicit material**.[23] Younger children again are at heightened risk. A global report that surveyed children in a range of countries, including in the global south, found that between **54% and 89% of children aged 9 to 11 years old reported having seen sexual content online in the past year**.[24]



Contact risks refer to children experiencing or being targeted by potentially harmful contact. For example, children risk being 'befriended' or groomed.[25] Potential offenders are able to contact children, meet them online, and build relationships directly with them pretending to be a child or presenting themselves under a fake social media, dating, or video gaming profile or using virtual identities.

[20] Risk categories included cyberbullying, sexual content, reputational risk, violent content, risky contact, and gaming disorder.
Park, Y., Gentile, D. A., Morgan, J., He, L., Allen, J. J., Jung, S. M., Chua, J., Koh, A. (2020) Child Online Safety Index Report, p.15. Available at: https://www.dqinstitute.org/wp-content/uploads/2020/02/2020-COSI-Findings-and-Methodology-Report.pdf (Accessed: 18 April 2022).

[21] 5Rights Foundation (2021) Pathways: how digital design puts children at risk. Available at: https://5rightsfoundation.com/uploads/Pathways-how-digital-design-puts-children-at-risk.pdf (Accessed 18 April 2022).

[22] UNICEF (2019) Global Kids Online Comparative report. Available at: https://www.unicef-irc.org/publications/pdf/GKO%20Main%20Report.pdf (Accessed: 18 April 2022).

[23] Madigan, S., Villani, V., Azzopardi, C., Laut, D., Smith, T., Temple, J. R., Browne, D., Dimitropoulos, G. (2018) 'The Prevalence of Unwanted Online Sexual Exposure and Solicitation Among Youth: A Meta-Analysis', The Journal of Adolescent Health, 63 (2), pp 133-141. https://doi.org/10.1016/j.jadohealth.2018.03.012.

[24] UNICEF (2019) *Global Kids Online Comparative report*.

[25] OSCE Office of the Special Representative and Co-ordinator for Combating Trafficking in Human Beings and Tech Against Trafficking (2020) Leveraging innovation to fight trafficking in human beings: A comprehensive analysis of technology tools. Available at: https://www.osce.org/files/f/documents/9/6/455206_1.pdf (Accessed: 18 April 2022).

> **Grooming**, or solicitation refers to *"the process of establishing/building a relationship with a child either in person or through the use of the Internet or other digital technologies to facilitate either online or offline sexual contact with that person."*[26]

Offenders typically build a trust relationship online with the child before requesting to meet offline or soliciting sexual content online. Studies show about **one in nine young people experience online sexual solicitation**.[27] Grooming mostly affects children aged 8 to 16 years old, with the majority between **11 and 13 years old**.[28] Girls are especially at risk of grooming.

"When I first created a Facebook account I couldn't understand how to use it and chat with others. A boy wanted my picture and I sent a picture of my hand. But later I understood [sending pictures] can be very harmful." (Focus group discussion, 15-year-old girl, Bangladesh)

"Everyone can lie on Internet and that can´t be controlled, it bothers me that people create fake profiles and ask you for naked pictures and those things." (Focus group discussion, Colombia)

The internet provides an expanding opportunity for potential offenders to contact a growing number of children. The many apps and online platforms provide more efficient ways to recruit, groom, and control victims. Offenders can easily create fake social media, dating, or video gaming profiles and use different fake virtual identities to reach out to children.[29] **Facebook-owned applications and Snapchat were the platforms used in over 70% of online grooming cases** where the platform used was known.[30] Once children are in communication with others, conduct and contract risks arise. Conduct risks are when children witness, participate in, or are the victims of potentially harmful conduct, such as sexual harassment. Importantly, children themselves may be actors here and may take actions that are potentially harmful to themselves or other children. An example of this might be non-consensual sharing or forwarding of sexual messages. In a study on potentially harmful online experiences among minors using social media, participants reported having online sexual interactions with someone they believed was an adult, with 15% of interactions on Snapchat, 13% on Instagram, 11% on WhatsApp, 10% on Facebook, and 10% on Messenger.[31]

Children become party to potentially harmful content. Around one third of CSAM reported for investigation is self generated by children themselves. Prevalence of this is sharply on the rise, as observed in the United Kingdom, where over 38,000 cases of **self-generated**[32] content were identified in the first quarter of 2021, more than double those found during the same period in 2020.[33] This may be because of the frequent practice of sexting among young people.

> **Sexting** refers to *"the creating, sharing, and forwarding of explicit sexual content, including sexually suggestive nude or nearly nude digital images and video. [...] It is often a consensual activity between peers, although research has shown that girls feel pressured or coerced into it more often than boys."*[34]

[26] ECPAT International (2016) *Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse*, p.51.

[27] Madigan, S., Villani, V., Azzopardi, C., Laut, D., Smith, T., Temple, J. R., Browne, D., Dimitropoulos, G. (2018) 'The Prevalence of Unwanted Online Sexual Exposure and Solicitation Among Youth: A Meta-Analysis'.

[28] NetClean (2019) NetClean Report 2018 – A Report about child sexual abuse. Available at: https://www.netclean.com/wp-content/uploads/2018/12/The-NetClean-Report-2018_Web.pdf (Accessed 25 April 2022).

[29] OSCE Office of the Special Representative and Co-ordinator for Combating Trafficking in Human Beings and T*ech Against Trafficking (2020) Leveraging innovation to fight trafficking in human beings: A comprehensive analysis of technology tools*.

[30] NSPCC (2021) Record high number of recorded grooming crimes lead to calls for stronger online safety legislation. 24 August. Available at: https://www.nspcc.org.uk/about-us/news-opinion/2021/online-grooming-record-high/ (Accessed 18 April 2022).

[31] Thorn (2021) *Responding to Online Threats: Minors' Perspectives on Disclosing, Reporting, and Blocking*. Available at: https://info.thorn.org/hubfs/Research/Responding to Online Threats_2021-Full-Report.pdf (Accessed: 18 April 2022).

[32] Internet Watch Foundation (2020) Face the facts – The Annual Report 2020. Available at: https://annualreport2020.iwf.org.uk/ (Accessed 18 April 2022).

[33] Internet Watch Foundation (2021) 'Appalling' rise of 'devious' criminals tricking children into sexually abusing themselves on camera. 31 May. Available at: https://www.iwf.org.uk/news-media/news/appalling-rise-of-devious-criminals-tricking-children- into-sexually-abusing-themselves-on-camera/ (Accessed 18 April 2022).

[34] ECPAT International (2016) *Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse*, p.44.

When sexting leads to abuse or exploitation, the Luxembourg Guidelines on Child Sexual Abuse stress that use of the term 'self-generated CSAM' should not result in blaming the child for what happens or in holding children criminally liable for the production of CSAM. Children are also self-generating such material in abusive contexts. Children who are groomed or coerced into exposing or abusing themselves on camera[35] might feel that they 'voluntarily' participated, however, offenders often dictate or orchestrate the act.[36]

The **nature of abuse** is also evolving: a survey of CSAM dark-net users indicates that **live-streamed CSAM** is becoming increasingly prevalent due to a number of factors, including the COVID-19 pandemic.[37] CSAM is widely exchanged, shared, and sold online, making the online dimension of this crime relevant.[38] The internet hosts platforms to attract clients and to deliver new types of sexual exploitation 'services', such as live streaming of child abuse. Skype, Snapchat and Facebook are the most commonly used online platforms for live-streamed child sexual abuse according to surveyed police officers.[39] Victims are procured and advertised through online platforms, apps, or escort/ sex work services sites.

Table 2. Escalation of reports of child sexual abuse online investigated by Internet Watch Foundation

335,558	361,000
1996 to 2011	2021 alone

Source: Internet Watch Foundation Annual Report 2021[40]

Organisations that investigate CSAM indicate that **2021 was the worst year on record**, with the number of reported cases escalating radically.[41] Numerous studies have shown that the dissemination of CSAM is problematic because it revictimises the children who are depicted in the imagery. In addition, it is also linked to further victimisation. CSAM has been shown to contribute to addiction, with many CSAM viewers having first encountered CSAM while under the age of 18. Secondly, viewing CSAM is linked to contacting children directly, with 42% of CSAM viewers having sought direct contact with children.[42]

Children who have been groomed and who have shared CSAM are further vulnerable to sexual extortion and offline exploitation, including trafficking.

> **Sexual extortion** of children (often called **sextortion**) is defined as *"the blackmailing of a person with the help of self-generated images of that person in order to extort sexual favours, money, or other benefits from her/him under the threat of sharing the material beyond the consent of the depicted person (e.g. posting images on social media)".*[43]

[35] Internet Watch Foundation (2020) *Face the facts – The Annual Report 2020*.

[36] INHOPE (2021) *What is self-generated CSAM*? Available at: https://www.inhope.org/EN/articles/what-is-self-generated-csam (Accessed 18 April 2022).

[37] Insoll, T., Ovaska, A. and Vaaranen-Valkonen, N. (2021) CSAM Users in the Dark Web: Protecting Children Through Prevention. Available at: https://drive.google.com/file/d/1EUBsU0A8XYw8QNUg3JKqemIRoLO9cYPt/view (Accessed 18 April 2022).

[38] ECPAT International (2016) *Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse,* p.40.

[39] NetClean (2019) NetClean Report 2019 – *A Report about child sexual abuse crime*, p.32. Available at: https://www.netclean.com/netclean-report-2019/ (Accessed 18 April 2022).

[40] Internet Watch Foundation (2022) *The Annual Report 2021*. Available at: https://annualreport2021.iwf.org.uk/trends/ (Accessed 6 May 2022).

[41] Internet Watch Foundation (2022) The Annual Report 2021; WeProtect Global Alliance (2021) *Global Threat Assessment 2021 – Working together to end the sexual abuse of children online*; NCMEC (2021) CyberTipline 2021 Report. Available at: https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata (Accessed 7 May 2022).

[42] Insoll, T., Ovaska, A. and Vaaranen-Valkonen, N. (2021) CSAM *Users in the Dark Web: Protecting Children Through Prevention*.

[43] ECPAT International (2016) *Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse*, p.27

Sexual extortion has been on the rise. Some national hotlines report drastically increased numbers in cases in comparison to pre-pandemic years.[44] Sexual extortion is most common with children aged 11 to 16 years old.[45] NetClean reports that, in their database, 78% of child victims of sexual extortion are girls.[46]

> **Child trafficking** is the *"recruitment and/or transport, transfer, harbouring, and receipt of a child by others with the intent of exploiting the child'. Most children are trafficked for sexual purposes".*

> *"There are adults that talk to teenage girls...they bother them. They tell them they have 'work' for them if they want work. [These are some of the] different things I have seen."*
> (Focus group discussion, 15-year-old girl, Bolivia)

As detailed above, children who are active online face several types of risks, and the likelihood and impact of harm is on the rise. In addition to the studies presented above, a multinational survey reported an increase in online grooming, sexual extortion, live-streamed CSAM, and self-generated CSAM.[48]

1.2.3 OSEC and CSAM are harmful and a source of 'never-ending' trauma for children

The online environment enables the rapid exchange of information, images, and videos. Images can be easily reshared in a few clicks, reaching thousands of users. Even temporary images and videos, or live-streamed material can be captured using screenshots and shared. OSEC survivors therefore not only endure the trauma of the actual abuse and exploitation, but are also further revictimised by the continuous circulation of the material, which brings a **permanent dimension to the abuse**.[49] CSAM survivors describe the chronic nature of this trauma related to the distribution of the material online as 'never ending'.[50] Victims of OSEC are *"living in fear of being recognised from photos and videos on the internet"*.[51] The fear of being recognised leads to major anxiety for victims combined with continuous revictimisation and the feeling of powerlessness.[52]

A survey of survivors across six countries shows that child victims experience fear and shame around their experience of OSEC. They fear being judged by their family, peers, or communities for their sexual conduct. Shame and fear may be heightened for LGBTQI+ children, who are scared of being outed and stigmatised for their sexual orientation.[53]

In addition, exposure to OSEC and victimisation have serious consequences for children. Research shows that exposure to sexual solicitations and explicit content lead to significant clinically diagnosable PTSD symptoms for child victims (e.g. hypervigilance or avoidance of thoughts or reminders of events).[54] In fact, children victims of OSEC have reported **similar levels of trauma symptoms** to child victims of **offline** sexual abuse.[55] Child victims of OSEC suffer from a wide range of psychological suffering, such as self-harming or suicidal behaviour, impaired relationships, difficulties at school, self-blame, and low self-esteem.[56]

[44] NCMEC (2022) Sextortion: The Hidden Pandemic. Available at: **https://www.missingkids.org/blog/2022/sextortion-the-hidden-pandemic** (Accessed 18 April 2022).

[45] NetClean (2019) *NetClean Report 2018 – A Report about child sexual abuse.*

[46] ibid

[47] ECPAT International (2016) *Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse*, p.73

[48] NetClean (2021) *NetClean Report – COVID-19 Impact 2020 – A Report about child sexual abuse crime.*

[49] NCMEC (2019) *Captured on film: survivors of child sex abuse imagery are stuck in a unique cycle of trauma.* Available at: **https://calio.org/wp-content/uploads/2020/03/Captured-on-Film-Survivors-of-Child-Sex-Abuse-Imagery-are-Stuck-in-a-Unique-Cycle-of-Trauma.pdf** (Accessed 18 April 2022).

[50] Canadian Centre for Child Protection Inc., *Survivors' Survey – Executive Summary 2017*. Available at: **https://www.protectchildren.ca/pdfs/C3P_SurvivorsSurveyExecutiveSummary2017_en.pdf** (Accessed 18 April 2022).

[51] Keller, M. and Dance, G. (2019) 'Images of Child Sexual Abuse. What Went Wrong?' *New York Times*, 28 September. Available at: **https://www.nytimes.com/interactive/2019/09/28/us/child-sex-abuse.html** (Accessed 18 April 2022).

[52] Joleby, M., Lunde, C., Landström, S. and Jonsson, L. S. (2020) '"All of Me Is Completely Different": Experiences and Consequences Among Victims of Technology-Assisted Child Sexual Abuse', Frontiers in Psychology, 11. **https://doi.org/10.3389/fpsyg.2020.606218**.

[53] ECPAT International and WeProtect Global Alliance (2022) *Child Sexual Abuse and Exploitation Online: Survivors Perspectives*. Available at: **https://www.weprotect.org/survivors-perspectives/** (Accessed 18 April 2022).

[54] McHugh, B. C., Wisniewski, P., Rosson, M. B. and Carroll, J. M. (2018) 'When social media traumatizes teens: The roles of online risk exposure, coping, and post-traumatic stress', Internet Research, 28(5), pp.1169-1188. **https://doi.org/10.1108/IntR-02-2017-0077**.

[55] Joleby, M., Lunde, C., Landström, S. and Jonsson, L. S. (2020) '"All of Me Is Completely Different": Experiences and Consequences Among Victims of Technology-Assisted Child Sexual Abuse'.

[56] ibid

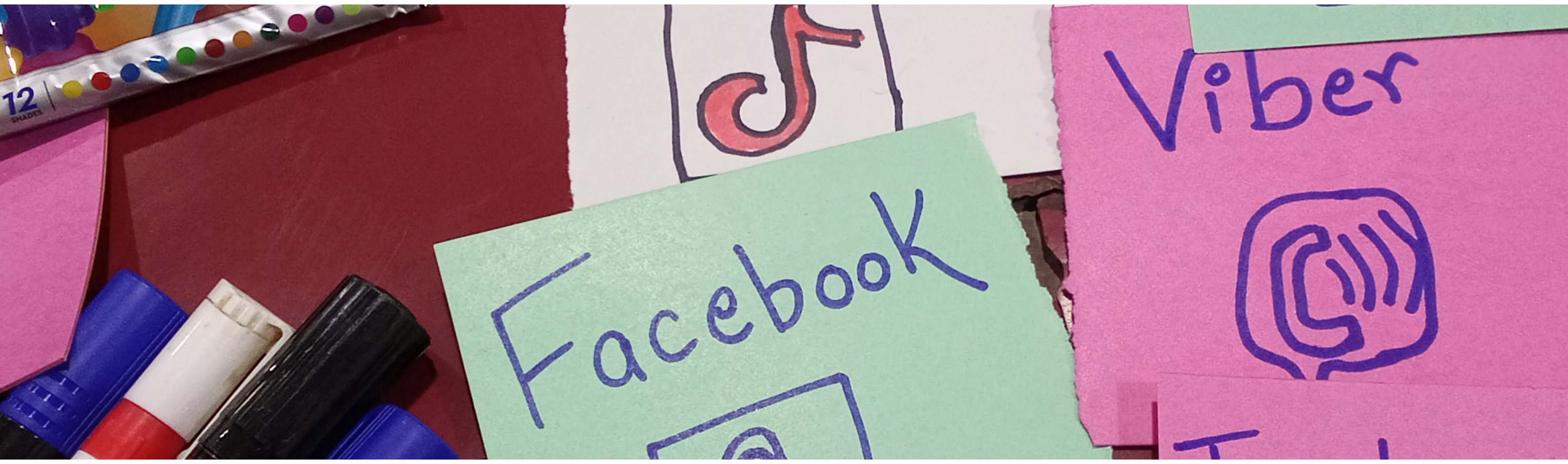
Research also reports that children who are frequently exposed to explicit content may become desensitised to its traumatic effects.[57] As mentioned above, exposure to sexually explicit content before the age of 18 may in fact lead an individual to seek out such material in adulthood. Suojellaan Lapsia's ReDirection survey of CSAM dark-net users revealed that 70% of 8,484 respondents had encountered CSAM before the age of 18, with 39% first exposed to it before the age of 13.[58] Other research has also confirmed that offenders involved in child sexual exploitation material offences have shown higher sexual interest in children than contact child sexual offenders.[59]

This brings to light the fact that children may themselves also commit OSEC against other children. Some evidence suggests that a substantial percentage of sexual harm against younger children is committed by people under the age of 18,[60] which shows the need for adequate prevention programmes for potential youth offenders. Similarly, research shows that those who resorted to sextortion were more likely to have been victims themselves.[61] Other studies have indicated a **high rate of OSEC victims among youth offenders**, while not specific to OSEC offences.[62]

It has been noted that sexual extortion has led to several cases of suicide of the child victim, as such violence often deeply affects child victims. In 2020, police officers surveyed in 39 countries reported an increase in sexual extortion, including cases ending in the suicide of the child victim.[63]

1.2.4 Digital environments are not adequately designed to keep children safe

The digital environment in which children are active influences their risk of OSEC. The way platforms are set up to facilitate contacts and the sharing of images and videos impacts the risks.[64] Many social media features designed to enhance communication put children using them at risk. 5Rights Foundation showed that **75% of the top 12 most popular social media platforms used AI to recommend child profiles to strangers**. The 'friend suggestions' expose children to the risk of grooming, as adults can list interests similar to children on their profiles. Other designs used by social media platforms that put children at risk include nudge techniques



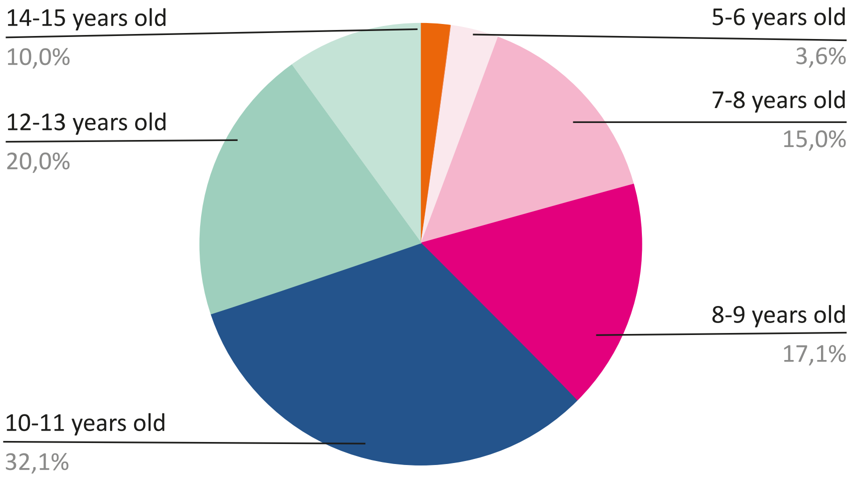
Source: Focus group in Nepal

[57] McHugh, B. C., Wisniewski, P., Rosson, M. B. and Carroll, J. M. (2018) 'When social media traumatizes teens: The roles of online risk exposure, coping, and post-traumatic stress'.
[58] Insoll, T., Ovaska, A. and Vaaranen-Valkonen, N. (2021) *CSAM Users in the Dark Web: Protecting Children Through Prevention*.
[59] Babchishin, K. M., Seto, M. C., Fazel, S. and Långström, N. (2019) 'Are There Early Risk Markers for Pedophilia? A Nationwide Case-Control Study of Child Sexual Exploitation Material Offenders', *The Journal of Sex Research*, 56(2), pp.203–212. https://doi.org/10.1080/00224499.2018.1492694.
[60] Letourneau, E. J., Schaeffer, C. M., Bradshaw, C. P. and Feder, K. A. (2017) 'Preventing the Onset of Child Sexual Abuse by Targeting Young Adolescents With Universal Prevention Programming', *Child Maltreatmen*t, 22(2), pp.100-111. https://doi.org/10.1177/1077559517692439.
[61] Patchin, J. W. and Hinduja, S. (2018) 'Sextortion Among Adolescents: Results From a National Survey of U.S. Youth', *Sexual Abuse*, 32(1), pp.30-51. https://doi.org/10.1177/1079063218800469.
[62] Cockbain, E. and Brayley, H. (2012) 'Child sexual exploitation and youth offending: A research note', *European Journal of Criminology,* 9(6), pp.689-700. https://doi.org/10.1177/1477370812453401.
[63] NetClean (2021) *NetClean Report – COVID-19 Impact 2020 – A Report about child sexual abuse crime*.
[64] Livingstone, S. and Smith, P.K. (2014) 'Annual Research Review: Harms experienced by child users of online and mobile technologies: the nature, prevalence and management of sexual and aggressive risks in the digital age', *The Journal of Child Psychology and Psychiatry*, 55(6), pp.635-54. https://doi.org/10.1111/jcpp.12197.

to connect with other users, visual popularity metrics encouraging young people to add strangers as 'friends' or 'followers'' and location tracking used by certain friend suggestion systems.[65] Furthermore, social media platforms are considered to be designed in a way that does not allow companies to easily distinguish between legitimate users and child offenders.[66] Lack of effective safety mechanisms on platforms widely used by children is an issue of concern.

Young children under the age of 13 are increasingly using the internet and social networking platforms that were not designed with their needs and risks in mind, and for which they may have limited preparation.[67] Despite age limits for young children on some social media platforms, various research shows that children under 13 join anyway. Out of the 141 children who participated in the focus groups for the current research, 98 (69%) were below the age of 12 when they first went on social media.

Graph 1: Age of first social media activity



Source: Focus groups in 10 countries across Asia, Europe and Latin America

Most social media platforms are not effectively identifying users who are younger than the recommended age for their platforms, nor are they protecting children from potential harm. Platforms predominantly rely on self-declaration (box-ticking or submitting a date of birth), but this has been found to be ineffective and easily circumvented by children.[68]

When asked whether it was okay to lie about their age online, children from the focus group discussions were rather divided: 54 (38%) agreed with the statement, while 52 (37%) stated they disagreed that it was fine to lie about their age and 35 (25%) did not know.

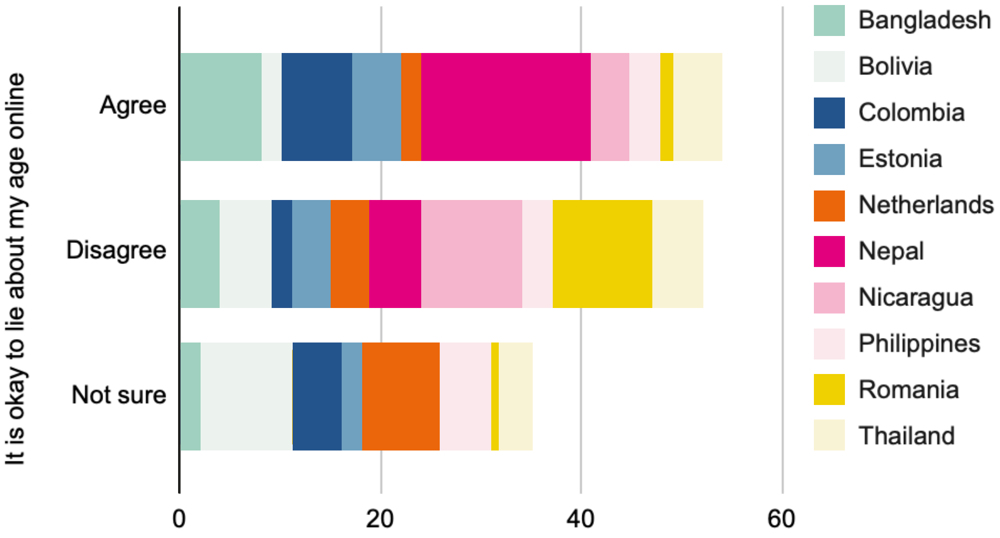
[65] 5Rights Foundation (n.d.) *Risky-by-Design*. Available at: https://www.riskyby.design/friend-suggestions (Accessed 18 April 2022).

[66] Salter, M. and Wong, T. (2021) *Research report – The impact of COVID-19 on the risk of online child sexual exploitation and the implications for child protection,* p.38. Available at: https://www.arts.unsw.edu.au/sites/default/files/documents/eSafety-OCSE-pandemic-report-salter-and-wong.pdf (Accessed 18 April 2022).

[67] UNICEF (2020) COVID-19 and its implications for protecting children online. Available at: https://www.unicef.org/sites/default/files/2020-04/COVID-19-and-Its-Implications-for-Protecting-Children-Online.pdf (Accessed 18 April 2022).

[68] Pasquale, L., Zippo, P., Curley, C., O'Neill, B and Mongiello, M. (2022) 'Digital Age of Consent and Age Verification: Can They Protect Children?', *IEEE Software*, 39(3), pp.50-57. https://doi.org/10.1109/MS.2020.3044872.

Graph 2: Lying about age online



Source: Focus groups in 10 countries across Asia, Europe and Latin America

Given the lack of safety-focused design online, and the difficulties with detecting and prosecuting perpetrators of OSEC, it is all the more important to focus on preventive measures. Research shows that service providers tend to **prioritise profit over user safety**,[69] exposing young users to the risk of sexual exploitation and CSAM (including self-generated material). While the impetus is often on reacting to harm that has already been done through the detection and removal of material, efforts need to be devoted to minimising exposure to risk in the first place. There is also a need to arm children with adequate and empowering tools for when they face risks online. **This calls for a stronger focus on the implementation of preventative measures to proactively reduce the number of OSEC cases** that are facilitated by social media platforms.

1.3 OBJECTIVES OF THE RESEARCH

This research aims to further understand child safety by design and how it can help better protect children online against OSEC. This includes identifying child safety by design solutions that have the potential to be effective and drafting proposed policy recommendations based on the findings. Broadly, the research seeks to achieve the following:

- **Problem definition**: Through a scoping literature review, we identified the scope of the problem and the role of safety by design. This led to the formulation of research questions that have been addressed by an in-depth literature review of academic sources related to child safety design, focusing on protecting children against OSEC, and understanding children's safety needs and current safety practices in use by the industry.
- **Identifying child safety by design solutions for OSEC**: Safety by design measures can be used to protect users from various risks. The research focuses specifically on OSEC and CSAM prevention. Child safety by design solutions were identified in the literature based on a set of criteria, including their potential drawbacks and privacy concerns, and children across Asia, Latin America and Europe were asked to contribute their ideas and views. Finally, a group of experts assessed the solutions and their effectiveness against OSEC.
- **Translating findings and solutions into EU policy recommendations:** A set of policy recommendations have been drawn from the research outcomes, which included the existing literature, focus group discussions with children, and an online workshop with experts. The recommendations call on the EU to adopt child safety by design requirements for social media platforms through its upcoming legislation.

[69] 5Rights Foundation (2021) *Pathways: how digital design puts children at risk*.

1.4 METHODOLOGY

The research process took place between November 2021 and April 2022 and consisted of the following steps:

1. **An initial scoping literature review** of the existing literature on the issue of child safety design. The goal was to identify the main issues related to child safety by design as well as gaps in research.

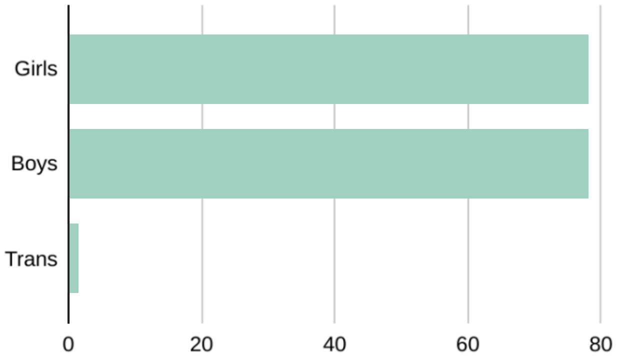
2. **An in-depth literature review** of academic sources related to child safety design. This review focused on 11 research questions specifically looking at how child safety design can effectively protect children against OSEC and CSAM.
   1: What safety designs are considered needed to keep children safe from OSEC and CSAM (including self-generated images and videos)?
   2: What safety by design measures are appropriate and effective depending on the age range of children?
   3: How can children's free expression be balanced with safety needs?
   4: How can we ensure safety measures are adapted to risks of OSEC/CSAM from adults as well as among children themselves?
   5: How can privacy be balanced with the need to protect children?
   6: Can age verification be used as a tool to prevent OSEC? If so, which age verification technologies should be considered?
   7: Should AI algorithms be deployed to protect children from OSEC and CSAM? What are the risks/drawbacks of using AI algorithms?
   8: Should parental controls be used to protect children from OSEC? How would that be tailored to a children's age?
   9: What factors prevent companies from implementing child safety by design? What is needed for companies to implement such design?
   10: What is the effectiveness of opt-in vs mandatory design?
   11: Which design/standard/practice should be regulated and made mandatory by EU policymakers to keep children safe online?

3. **Focus group discussions (FGDs) with children in DtZ countries and selected EU countries** to gather information on how children feel about child safety by design, their online presence, and whether more rules should exist to protect them. The focus group discussions also served as an opportunity for children to generate ideas about safe design and ensured that children's voices are included in the research and in EU advocacy. Focus groups took place in the following 10 countries:
   - **Asia:** Nepal, Bangladesh, Philippines, Thailand
   - **LATAM:** Bolivia, Colombia, Nicaragua
   - **Europe:** Estonia, Romania, the Netherlands

   The focus groups were carried out by the following organisations: **Bangladesh**: Terre des Hommes Netherlands Bangladesh office and ASK; **Bolivia**: Vuela Libra; **Colombia**: Fundación Renacer; **Estonia**: OÜ Cyberlist; **Nepal**: Plan International Nepal office, Terre des Hommes Netherlands Nepal office with CWIN and WYESHR; **Netherlands**: Stichting Alexander; **Nicaragua**: Tesis Asociación; **the Philippines**: Terre des Hommes Netherlands Philippines office; **Romania**: Terre des Hommes Lausanne Romania office; **Thailand**: ECPAT Foundation Thailand.

   A total of 141 children, aged 11 to 16 years old, participated in the focus groups across the countries. A diverse group of children participated in the focus groups across the countries. Some groups included very at risk children from (very) low economic background while others included middle class background children with lower risk of OSEC. The groups also included runaways, victims of sexual violence, some had little parental supervision, tense relationships with their parents, history of violence /prostitution in the family, single parent household, children with substance abuse issues, migrant status, ethnically marginalised group background, mental disabilities, orphan status and addiction to being online. All children were active online. Only one child identified as trans and another as gay. A little over half the child participants were girls (78), with a good representation of boys (62) and one trans participant.

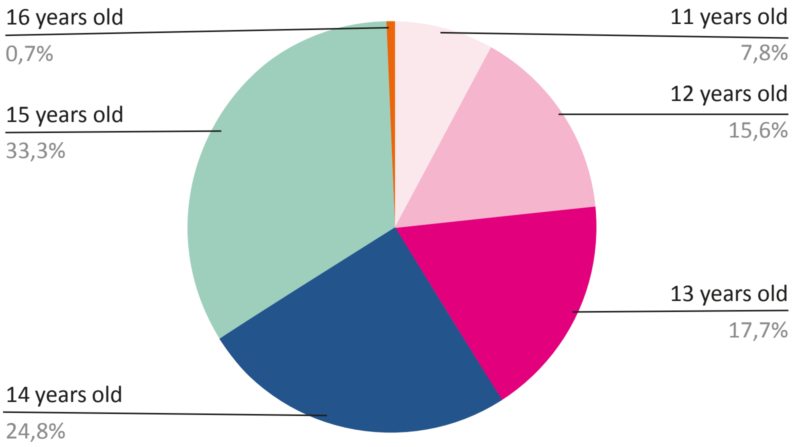
Graph 3. Gender of focus group participants



Source: Focus groups in 10 countries across Asia, Europe and Latin America

The vast majority of child participants were 13 years old and above (77%), while 23% were between 11 and 12 years old.

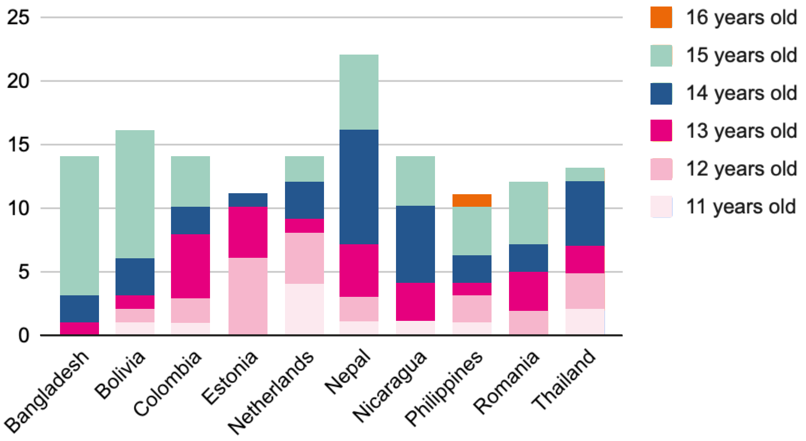
Graph 4. Age of focus group participants



Source: Focus groups in 10 countries across Asia, Europe and Latin America

The age groups varied across countries, with some countries having mostly teenagers of 14 to 15 years old, and others having younger teenagers.

Graph 5: Age of FGD child participants across countries



Source: Focus groups in 10 countries across Asia, Europe and Latin America

4. **Mapping of solutions** discussed in the literature, as well as those raised by children during FGDs. A resulting set of solutions was identified. A group of senior experts, selected on the basis of their expertise and representing a variety of practitioners and academic experts, were invited to identify and evaluate child safety by design measures against the following criteria:
   - Effectiveness in preventing OSEC and CSAM, reducing risks of inadequate contacts, and reducing inappropriate exchange of images, videos, or chats.
   - Likelihood of being circumvented by children.
   - Respectful of user privacy, in line with the GDPR and minimising privacy risks.
   - Respectful of children's rights, balancing their right to be heard and their right to participate in society and in decisions affecting them, as well as being in the best interests of the child.
   - Allowing, and not preventing children from accessing the digital environment in a safe way.

5. **Online expert workshop validating EU policy recommendations.** A long list of recommendations was developed through the literature review, focus groups with children, and the solution mapping. From this list, an initial shorter list was discussed in an online workshop on 17 March 2022. Twenty senior experts participated in the workshop, providing feedback on the recommendations. From the feedback, a final list of recommendations was prepared and is presented in this paper.

1.4.1 Limitations of the research

The research findings have to be seen in the light of the limitations of this research. A first limitation is that child safety by design is a **very broad topic**. While this research has aimed to be as complete as possible, there are still areas within child safety by design that have not been included. The detection of OSEC, for instance, is something that is not specifically discussed in this research. Another limitation is that where it was tried to have as much global literature as possible, the reality is that many of the sources were **Western-European focused**. The **lack of transparency of** online platforms also hampered the ability to be complete in this sense, as information about what measures platforms currently use to keep children safe is often not available to the public. This limitation of incompleteness also goes for the solutions. There are many more solutions that will aid to enhance the safety of children online, but it is impossible to extensively study all of them. The solutions and recommendations in this research should therefore not be seen as an exhaustive list, but rather as pieces to the puzzle that will contribute to a better safety design for children.

Another set of limitations stems from the focus groups. In total, 141 children participated in focus groups in ten countries. While this is a substantive group, it is **not a number that provides statistically significant outcomes.** Moreover, the **characteristics of the children in the groups varied** a lot. In some groups, children at risk of (O)SEC were specifically targeted to be in the focus group discussions, whereas in other groups, this was not the case. Conducting the focus groups in ten different countries across three continents provided the research with many diverse insights, but also makes it harder to compare data between groups.

Despite these limitations, this research was able to highlight many important issues that will help both the industry and the EU to strengthen child safety by design. Informed by an extensive literature review, this report provides an overview of many factors that are important to keep in mind when designing online platforms, stressing the complicated field that online platforms find themselves in.

2. Key considerations for child online safety and platforms’ safety measures

KEY MESSAGES

Gender and intersectionality

- OSEC is a form of **gender-based violence**. Gender norms shape the manifestation of the violence and the risks of victimisation, with adolescent girls most at risk
- **Gender norms** of submissive behaviour for girls and dominance and assertiveness for boys affect their online behaviours and risks of OSEC
- Adherence to the hegemonic male gender role can cause masculine gender role stress in male adolescents, which is a predictor of violence against women and girls
- Children with **disabilities**, children who question their sexuality or self-identify as **LGBIQI+** or as a **racial or ethnic minority** are also more likely to experience sexual harm online
- **Risk factors** for OSEC include: low self-esteem, mental health issues, disruptive or dysfunctional family dynamics, struggles with social interactions, and impulsivity

Children’s rights

- Children have a number of **rights** when active online, including: proactive **prevention** of discrimination and sexual exploitation, equal access, non-discriminatory design, and that their **best interests** are a primary consideration in the provision, regulation, design, management, and use of the digital environment

Understanding children’s online behaviour

- As younger and older children have different needs, design should be tailored in such a way that children can **increasingly have more responsibility and freedom** as they get older
- Teenagers go through **many changes** during adolescence, making them vulnerable due to increased need for social and sexual interactions, but slower cognitive development. Online predators can take advantage of this, often making a **vulnerability assessment** to select their target
- The **fear of stranger danger** is overestimated, as young people are more often victimised by people they know, including children themselves

Safety measures used by platforms

- Social media apps increasingly use **artificial intelligence** (AI) to actively search for inappropriate content and risky behaviour, but only Meta intervenes by sending warning messages
- Meta, TikTok and YouTube take a more **proactive approach** towards keeping children safe online and removing inappropriate content, while Snapchat is more reactive

Challenges

- Ensuring safety online for children requires keeping a **balance** between their protection and the full realisation of their right to **freedom of expression** and right to **privacy**
- Any intrusion into the right to freedom of expression and right to privacy must be **lawful, necessary** and **proportional**
- When making design decisions, companies tend to prioritise profit over safety, including the use of dark patterns that are not privacy friendly
- Designers face a number of **challenges** such as the need to make apps financially profitable, the use of third-party libraries, and a lack of awareness/guidelines on ensuring child safety

In this Section, we aim to provide an overview of safety by design, including how it is currently in use and the challenges it faces. In order to provide a comprehensive understanding of safety by design, the Section includes some key considerations to enable online designs to be effective in preventing online harm such as OSEC.

2.1 KEY CONSIDERATIONS FOR ONLINE CHILD SAFETY BY DESIGN SOLUTIONS

Before looking at the safety features currently in use in Section 2.2, this Section provides some key considerations for child safety by design solutions. Intersecting factors that exacerbate risks of OSEC are introduced, with a clear focus on OSEC as a form of gender-based violence. Thereafter, the rights of children in the digital environment are outlined, as solutions need to balance child protection with children's rights to participation in the online environment.

2.1.1 Intersectionality and compounding vulnerability to OSEC

OSEC does not happen in a vacuum. Its manifestations follow certain patterns. OSEC affects children of all ages, gender and backgrounds. Yet, certain groups are more at risk. Personality and behaviour influence the way children use the internet and their likelihood to be exposed to harm such as OSEC.[70] However, gender, age, sexuality, and disability require additional attention. Intersectionality provides a lens to ensure that all the relevant factors are acknowledged and brought into the response to a particular group.

> **Intersectionality** *is a theoretical framework that considers overlapping or intersecting factors that can be empowering or oppressive,*[71] *and that ultimately puts some people at a disadvantage.*

When referring to children as a group, diversity and potential intersections need to be accounted for. There are a number of intersecting factors that compound the risk of exposure to OSEC for some children. Various personal characteristics or circumstances can also be protective factors that minimise risks for some. Key evidence on the influence of gender and age are presented in the sections below.

WHICH CHILDREN ARE MOST AT RISK?	
Sex	Girls are more likely to be groomed, exploited, and featured in CSAM. Girls are consequently more likely to have negative sexting experiences. Girls are more likely to be blamed for their victimisation. Boys are less likely to report victimisation. Boys are more at risk of sexual abuse by someone outside the family than girls. Boys are more likely to be exposed to sexual material and pornography.
Age	Children are more vulnerable at the onset of puberty (11-13 years for girls and 14-15 years for boys). Prevalence is growing most rapidly among children under the age of 13. Younger children are more at risk of online abuse from known family members, peers, and friends. Adolescents are more likely to engage in risky online sexual activities including unsafe sexting and seeking violent sexual content, including CSAM.

[70] Livingstone, S. and Smith, P.K. (2014) 'Annual Research Review: Harms experienced by child users of online and mobile technologies: the nature, prevalence and management of sexual and aggressive risks in the digital age', ***The Journal of Child Psychology and Psychiatry***, 55(6), pp.635-54. **https://doi.org/10.1111/jcpp.12197**.

[71] Crenshaw, K. (1989) 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', ***University of Chicago Legal Forum,*** 1989(1). Available at: **https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf** (Accessed 7 May 2022).

Which children are most at risk?	
Disability	Children with intellectual impairments are more likely to be groomed. Children with disabilities are more likely to experience online sexual harm by an adult they know.
Sexuality	Children who are LGBTQI+ or who are questioning their sexuality are more likely to experience online sexual harm.

2.1.2 Gender is a determinant factor in OSEC

OSEC is a form of gender-based violence. This means that gender norms shape the manifestation of the violence and the risks of victimisation. Studies show that adolescents, in particular adolescent girls, are more at risk of grooming and other online harms.[72]

When addressing online risks, it is important to understand the patterns of children's social media use. Research has found a link between time spent online and increased likelihood of being a victim of harm. Yet, time spent online is not a sole determinant and must be combined with risk-taking behaviours and other risk factors. Research conducted by Australia's eSafety Commissioner (2017) among children and teenagers shows that girls are more likely to be using social media (especially Instagram and Snapchat) than boys.[73] Data from the EU Kids Online 2020 survey indicates that girls (57%) are more likely to visit social media sites daily than boys (51%), while boys are twice as likely to play online games than girls.[74] Increased use of social media translates to an increased risk of exposure to online harms.

Girls are more likely to be victims of OSEC than boys. Worldwide, one in five girls and one in 13 boys have been sexually exploited or abused before reaching the age of 18.[75] Depending on the form of violence, the difference can be more or less important. Of child victims depicted in CSAM reported to INHOPE hotlines, 93% were girls, 5% were boys and 2% included both genders.[76] A WeProtect survey reports that girls (24%) are twice as likely than boys (12%) to have someone they did not know ask them to do something sexually explicit online that made them feel uncomfortable or that they did not want to do.[77] UK data shows that girls were victims in 83% of the grooming cases reported between 2017 and 2021.[78]

[72] Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming', *Aggression and Violent Behavior,* 18, pp.135-146. https://doi.org/10.1016/j.avb.2012.11.008; Jonsson, L. S., Fredlund, C., Priebe, G., Wadsby, M. and Svedin, C. G. (2019) 'Online sexual abuse of adolescents by a perpetrator met online: a cross-sectional study', *Child and Adolescent Psychiatry and Mental Health*, 13, 32. https://doi.org/10.1186/s13034-019-0292-1; Cooper, K., Quayle, E., Jonsson, L. and Svedin, C. G. (2016) 'Adolescents and self-taken sexual images: A review of the literature', *Computers in Human Behavior*, 55, part B, pp.706-716. https://doi.org/10.1016/j.chb.2015.10.003.

[73] eSafety Commissioner (n.d.) *Young people and social media usage*. Available at: https://www.esafety.gov.au/research/youth-digital-dangers/social-media-usage (Accessed 18 April 2022).

[74] Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S., and Hasebrink, U. (2020) *EU Kids Online 2020: Survey results from 19 countries*. https://doi.org/10.21953/lse.47fdeqj01ofo.

[75] UNICEF (2021) *Ending online child sexual exploitation and abuse: Lessons learned and promising practices in low- and middle-income countries*. Available at: https://www.unicef.org/media/113731/file/Ending%20Online%20Sexual%20Exploitation%20and%20Abuse.pdf (Accessed 18 April 2022).

[76] INHOPE (2021) *Annual Report 2020*. Available at: https://inhope.org/media/pages/the-facts/download-our-whitepapers/annual-report/bb4dd3cdc3-1628156678/inhope-annual-report-2020.pdf (Accessed 18 April 2022).

[77] Economist Impact and WeProtect Global Alliance (2022) *Estimates of childhood exposure to online sexual harms and their risk factors*. Available at: https://www.weprotect.org/economist-impact-global-survey/#report (Accessed 18 April 2022).

[78] NSPCC (2021) *New figures reveal four in five victims of online grooming crimes are girls*. Available at: https://www.nspcc.org.uk/about-us/news-opinion/2021/online-grooming-crimes-girls/ (Accessed 18 April 2022).

Girls are consequently more likely than boys to have negative sexting experiences, be harassed by sexts from strangers, and be pressured to send sexts.[79] Laura Bates' Everyday Sexism Project showed that young girls report feeling pressure to sext and send sexualised material to boys. The Project also uncovered the impact of exposure to pornography at a young age, a lack of understanding of consent in the online environment (including of further disseminating images without consent), and of what behaviour falls under sexual violence.[80]

Research confirms that adolescent girls feel conflicted between refusing and pleasing their partner. When requests for unwanted sexual material come from peers or known adults, girls fear being rejected or shamed by them. Boys are shown to shame, block, or show hostility to girls who decline requests.[81]

In addition to cisgendered heterosexual girls, **social norms about gender identity and sexual orientation contribute to violence against LGBTQI+ children.**[82] Children who question their sexuality[83] are more likely to experience sexual violence. The Economist Impact global survey reports that 65% of LGBTQI+ respondents experienced online sexual harm as compared to 46% of non-LGBTQI+ respondents.[84] A study found that self-reported homosexuality or bisexuality was a strong risk factor for being approached sexually online.[85]

Boys are less likely to be victims of OSEC, and have different support needs to girls and LGBTQI+ children. While it is estimated **that OSEC against boys is underreported**, studies report that boys are more at risk of sexual abuse by someone outside the family than girls. Boys are more likely to be exposed to sexual material and pornography.[86] Gender norms also affect boys, where toxic masculinity and the requirement of male dominance is perceived as incompatible with victimisation. Boys victims can be shamed and emasculated for their experience with online sexual abuse. The strong unequal gender attitudes and cultural and social norms around masculinity create barriers for boys to report victimisation. Boys may be perceived to be lucky to receive sexual attention from girls or women even if unwanted and non-consensual.[87] Risk factors for the victimisation of boys include *"diverse sexual orientation or gender identity, homophobia/stigma, lack of awareness that boys can be sexually abused, perception that abuse of boys is not serious or harmful [...] and "macho" male image".*[88]

While victims tend to be mostly girls, **offenders are mostly men**. Boys and girls are more likely to be abused by a male offender. This was also recognised in the focus group discussion. In response to the question of what they disliked about social media, the children mentioned:

"There are older men harassing girls and boys." (Focus group discussions, girl, Nicaragua)

"Some strange (older) men also try to interact. I block them."
(Focus group discussions, 12-year-old-girl, Estonia)

"Men are nasty and rude. I do not like being talked to by 40 year old men who are rude and say nasty things."
(Focus group discussions, Colombia)

[79] Burén, J. and Lunde, C. (2018) 'Sexting Among Adolescents: A Nuanced and Gendered Online Challenge for Young People', ***Computers in Human Behavior,*** 85, pp.210-217. https://doi.org/10.1016/j.chb.2018.02.003.

[80] Bates, L. (2015) ***Everyday Sexism***. Simon & Schuster UK.

[81] Mishna, F., Milne, E., Cook, C., Slane, A. and Ringrose, J. (2021) 'Unsolicited Sexts and Unwanted Requests for Sexts: Reflecting on the Online Sexual Harassment of Youth', ***Youth & Society***. https://doi.org/10.1177/0044118X211058226.

[82] UNICEF (2020) ***Gender Dimensions of Violence Against Children and Adolescents***. Available at: https://www.unicef.org/media/92376/file/Child-Protection-Gender-Dimensions-of-VACAG-2021.pdf (Accessed 18 April 2022).

[83] Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[84] Economist Impact and WeProtect (2022) ***Estimates of childhood exposure to online sexual harms and their risk factors***.

[85] Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[86] Livingstone, S. and Helsper, E. (2010) 'Balancing opportunities and risks in teenagers' use of the internet: the role of online skills and internet self-efficacy'. ***New Media & Society***. 12(2), pp.309-329. https://doi.org/10.1177/1461444809342697; Cabello-Hutt, T., Cabello, P. and Claro, M. (2018) 'Online opportunities and risks for children and adolescents: The role of digital skills, age, gender and parental mediation in Brazil'. ***New Media & Society***. 20(7), pp.2411-2431. https://doi.org/10.1177/1461444817724168.

[87] Bradbury, P. and Martellozzo, E. (2021) ''Lucky Boy!'; Public Perceptions of Child Sexual Offending Committed by Women', ***Journal of Victimology and Victim Justice***. https://doi.org/10.1177/25166069211060091.

[88] ECPAT International (2021) Global Boys Initiative: ***A global review of existing literature on the sexual exploitation of boys***. Available at: https://ecpat.org/wp-content/uploads/2021/09/Global-Boys-Initiative-Literature-Review-ECPAT-International-2021.pdf (Accessed 7 May 2022).

Although women make up a low proportion of offenders, the rate of female offenders that sexually abuse boys is slightly higher than the rate of female offenders victimising girls.[89] **Boys and men are more likely to ask for sexts than females**.[90] Surveyed police officers report that aside from being mostly male, "there are no other typical attributes such as age, profession, family situation or otherwise" of child sexual abuse offenders.[91]

Unequal gender norms, heteronormative male dominance behaviour, and men violence normalised by patriarchal societal structures and systems contribute to violence against children, including OSEC. Patriarchal societies encourage men's power and control over women and children. This includes power over bodily integrity, which reinforces subordination. Boys and men who embrace unequal gender norms and ideals of male dominance are more likely perpetrate violence against women and children.[92]

Research also shows that girls and boys are socialised into different sexual behaviours and are held to different standards as to their sexual behaviour online or offline. Boys are expected or even praised for sexual behaviour, while girls are blamed and disapproved of for the same behaviour. Sexting among male adolescent tends to be perceived as normal, leading to status gain among peers, while female adolescents who sext receive more negative reactions and tend to be shamed for the same behaviour.[93]

> *"T[t]raditionally, men/boys are expected to be sexually active, dominant, and the initiator of (hetero) sexual activity, whereas women/girls are expected to be sexually reactive, submissive, and passive. Moreover, traditionally men are granted more sexual freedom than women. As a consequence, women and men are treated differently when they show the same sexual behaviors."*[94]

When violence occurs, women and children are often blamed for violating expected submissive behaviour norms. Teenage girls are often blamed for 'nudes' of them being shared, when the girl herself did not consent to the forwarding and distribution of a sext message she sent to a boyfriend. **Victim-blaming** attitudes and prioritisation of the offender's reputation is widespread and helps validate violence as a legitimate form of social control.[95]

The difference in socialisation and **internalised sexism** impacts risk exposure to online harm and can also lead adolescents to offending. Internalised misogyny has a key role in the perpetuation of online violence by male teenagers against female teenagers. A recent study reports that the 'justification of male dominance and violence' and risky sexual behaviours online were key factors contributing to male adolescent sexual harassment of girls online, while low self-esteem and risky sexual behaviours online were factors contributing to victimisation of female adolescents.[96]

[89] Cockbain, E., Ashby, M. and Brayley, H. (2015) 'Immaterial boys? A large-scale exploration of gender-based differences in child sexual exploitation service users', *Sexual Abuse*, p.5. https://doi.org/10.1177/1079063215616817.

[90] Walrave, M., Heirman, W. and Hallam, L. (2013) 'Under pressure to sext? Applying the theory of planned behaviour to adolescent sexting'. *Behaviour & Information Technology,* 33(1), pp.86-98. https://doi.org/10.1080/0144929X.2013.837099.

[91] NetClean (2019) NetClean Report 2018 – *A Report about child sexual abuse*. Available at: https://www.netclean.com/wp-content/uploads/2018/12/The-NetClean-Report-2018_Web.pdf (Accessed 25 April 2022).

[92] UNICEF (2020) *Gender Dimensions of Violence Against Children and Adolescents*.

[93] Endendijk, J. J., Deković, M., Vossen, H., van Baar, A. L. and Reitz, E. (2022) 'Sexual Double Standards: Contributions of Sexual Socialization by Parents, Peers, and the Media'. *Archives of Sexual Behavior*, 51, pp.1721-1740. https://doi.org/10.1007/s10508-021-02088-4; Walrave, M., Heirman, W. and Hallam, L. (2013) 'Under pressure to sext? Applying the theory of planned behaviour to adolescent sexting'.

[94] Endendijk, J. J., Deković, M., Vossen, H., van Baar, A. L. and Reitz, E. (2022) 'Sexual Double Standards: Contributions of Sexual Socialization by Parents, Peers, and the Media'.

[95] Namy, S., Carlson, C., O'Hara, K., Nakuti, J., Bukuluki, P., Lwanyaaga, J., Namakula, S., Nanyunja, B., Wainberg, M., Naker, D. and Michau, L. (2017) 'Towards a feminist understanding of intersecting violence against women and children in the family', *Social Science & Medicine*, Vol. 184, pp.40-48. https://doi.org/10.1016/j.socscimed.2017.04.042; UNICEF (2020) *Gender Dimensions of Violence Against Children and Adolescents*.

[96] Díaz-Aguado, M. J. and Martínez-Arias, R. (2022) 'Types of Male Adolescent Violence Against Women in Three Contexts: Dating Violence Offline, Dating Violence Online, and Sexual Harassment Online Outside a Relationship', *Frontiers in Psychology,* 13, 850897. https://doi.org/10.3389/fpsyg.2022.850897.

Source: Focus group in Nepal

Masculine Gender Role Stress (MGRS) is a term that encapsulates the emotional stress experienced for not adhering to or for violating masculine gender norms. MGRS can be the result of physical inadequacy, emotional inexpressiveness, subordination to women/being outperformed by women, and intellectual inferiority. MGRS is shown to be a key factor in violence against women. For instance, research indicates that experiences of MGRS related to situations of subordination to women at the beginning of adolescence is a better predictor of male adolescent dating violence than low self-esteem.[97]

Similarly, adherence to hegemonic male gender role beliefs, such as 'the expectation that men are emotionally and physically tough and willing to be aggressive', and 'the belief that men should not engage in stereotypically feminine activities', are positively associated with men's hostility toward women.[98]

Education programmes on sexuality as well as sexual and reproductive health rights tend not to be sufficiently comprehensive or are simply lacking in terms of consent, exploring sexuality online, sexist attitudes, and gender sensitivity. The stigma of such topics often means that adolescents learn from their peers, media, online platforms, or pornography, with a high risk of perpetuating sexist and violent behaviours they have internalised.[99]

2.1.3 Age and child development affect online risks

Age is another factor that intersects with gender to contribute to CSE vulnerability.[100] Age affects the way children use the internet, the time they spend online, and the way they use social media. Research indicates that **different ages imply different needs and risks**.[101] Age, development, and maturity affect the risks children face and how they are able to perceive and respond to them. Younger children are more vulnerable when exposed to sexual encounters online compared to their older counterparts or adults. This is because they have not yet undergone the physical, social and cognitive development of adolescence.[102]

Apps should therefore not be designed for children as a general group, but design should be tailored to different age groups, taking gender into account. This could be achieved by having features that the child can customise themself. The app could make suggestions for what would be age-appropriate for the user in question.[103]

[97] Merino, E., Díaz-Aguado, M. J., Falcón, L. and Martínez-Arias, R. (2021) 'Masculine Gender Role Stress as a Mediator of the Relationship Between Justification of Dominance and Aggression and Male Adolescent Dating Violence Against Women', *Psicothema*, 33(2), pp.206-213. https://doi.org/10.7334/psicothema2020.275.

[98] Gallagher, K. E. and Parrott, D. J.(2011) 'What accounts for men's hostile attitudes toward women? The influence of hegemonic male role norms and masculine gender role stress', *Violence Against Women*, 17(5), pp.568-583. https://doi.org/10.1177/1077801211407296.

[99] Sharma, M.K., Anand, N., Kumar, K., Lenin Singh, R., Thakur, P.C., Mondal, I., Kohli, T. (2021) 'Constructing the understanding of teenagers deviant use of cyberspace'. *International Journal of Social Psychiatry*. 67(8), pp.1068-1071. https://doi.org/10.1177/0020764020975791.

[100] Laird, J., Klettke, B., Hall, K., Clancy, E. and Hallford, D. (2020) 'Demographic and Psychosocial Factors Associated With Child Sexual Exploitation – A Systematic Review and Meta-analysis', *Jama Network Open*, 3(9). https://doi.org/10.1001/jamanetworkopen.2020.17682.

[101] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls*. Available at: http://eprints.lse.ac.uk/112559/ (Accessed 22 April 2022); Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewski, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online', *18th ACM International Conference on Interaction Design and Children*, Boise, United States, 12-15 June. ACM, New York, United States, pp.394-406. https://doi.org/10.1145/3311927.3323133.

[102] Burén, J. and Lunde, C. (2018) 'Sexting Among Adolescents: A Nuanced and Gendered Online Challenge for Young People'.

[103] Ghosh, A. K., Hughes, C. E. and Wisniewski, P. (2020) 'Circle of Trust: A New Approach to Mobile Online Safety for Families', *2020 CHI Conference on Human Factors in Computing Systems*. Honolulu, United States, 25-30 April. ACM, New York, United States, pp.1-14. https://doi.org/10.1145/3313831.3376747.

CHILDREN UNDER 13

Children use the online world for many different purposes, such as entertainment and socialising, but also for education as they use the internet for learning and school.[104] The use and nature of the internet and social media changes as children grow up. A survey of internet use at home found that the **use of the internet strongly increases with age**.[105] The internet provides opportunities for children to learn about various topics, but younger children need more guidance from adults in how to safely use the internet.

While adolescent online behaviour and OSEC risks are well described, there is **less literature available that focuses on younger children**. This is a clear gap, since available data on internet use and OSEC prevalence suggests that children under the age of 13 are currently an expanding user group that encounters risk online. Cyber Safe Kids reported that most (84%) 8- to 12-year-olds were using social media and messaging apps in 2020.[106] Furthermore, European data, also from 2020, found that in some countries, 45% of children aged 9 to 11 visited social networking sites daily.[107] Children aged 11 and 13 years old are most susceptible to grooming,[108] linked to the increasing trend of CSAM cases featuring children under 13.[109]

While the common conceptualisation of grooming often involves a stranger predator, this is not always the case. Studies hint at younger children being more at risk of OSEC and online abuse from **known family members, peers, and friends than from strangers**. This is because young children have different internet usage and if they have social media profiles, their privacy settings are at the highest, allowing only close family members and friends to interact with the child, while adolescents have more freedom as to their use of social media.[110]

Preadolescent children require **more guidance and supervision** than their teenage counterparts. Younger children are lacking certain information that is needed for them to fully comprehend online risks. Research by Badillo-Urquiola et al. (2019) found that younger children between the ages of 7 to 11 years old firstly were able to recognise risky situations and secondly, have a need for autonomy.[111] This underlines the importance of taking children's wishes into account when designing parental control features.[112] For younger children, several platforms have also launched a child-friendly version, such as YouTube Kids, where there is more monitoring and safe content.[113]

[104] Nash, V., Adler, J. R., Horvath, M. A. H., Livingstone, S., Marston, C., Owen, G. and Wright, J. (2015) *Identifying the routes by which children view pornography online: implications for future policy-makers seeking to limit viewing*. Available at: http://eprints.lse.ac.uk/65450/1/__lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_Livingstone%2C%20S_Identifying%20the%20routes_Livingstone_Identifying%20the%20routes_2016.pdf (Accessed 22 April 2022).

[105] Livingstone, S. and Smith, P.K. (2014) 'Annual Research Review: Harms experienced by child users of online and mobile technologies: the nature, prevalence and management of sexual and aggressive risks in the digital age'; Nash, V., Adler, J. R., Horvath, M. A. H., Livingstone, S., Marston, C., Owen, G. and Wright, J. (2015) *Identifying the routes by which children view pornography online: implications for future policy-makers seeking to limit viewing.*

[106] Cyber Safe Kids (2021) *Annual Report 2020*. Available at: https://www.cybersafekids.ie/wp-content/uploads/2021/09/CSK_Annual_Report_2020_web.pdf (Accessed 18 April 2022).

[107] Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S., and Hasebrink, U. (2020) EU Kids Online 2020: *Survey results from 19 countries*.

[108] NetClean (2019) *NetClean Report 2018*.

[109] INHOPE (2021) *Annual Report 2020*.

[110] Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewski, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online'; Smith, P. K., Thompson, F. and Davidson, J. (2014) 'Cyber safety for adolescent girls: bullying, harassment, sexting, pornography, and solicitation', *Current Opinion in Obstetrics and Gynecology*, 26(5), pp.360-365. https://doi.org/10.1097/GCO.0000000000000106.

[111] UNICEF (2017) *The State of the World's Children 2017 – Children in a Digital World*. Available at: https://www.unicef.org/media/48601/file (Accessed 25 April 2022).

[112] Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewki, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online'.

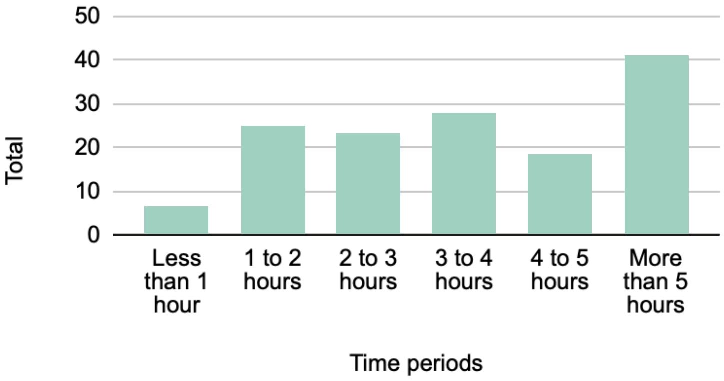
[113] Nash, V., Adler, J. R., Horvath, M. A. H., Livingstone, S., Marston, C., Owen, G. and Wright, J. (2015) *Identifying the routes by which children view pornography online: implications for future policy-makers seeking to limit viewing*.

ADOLESCENTS (13 TO 17 YEARS OLD)

The EU Kids Online 2020 survey, conducted across 19 countries between autumn 2017 and summer 2019, reported a significant increase in the number of children using smartphones and the amount of time they spend online. The survey showed that children spent twice as much time online in 2017-2019 compared to the results of the 2010 EU Kids Online Survey. Considering this was before the COVID-19 pandemic, it can reasonably be expected that the further upturn in screen-time is substantial. This is also related to the increase in the use of smartphones, enabling children to be connected to the internet almost constantly. Older children tend to spend more time online than younger children. Children of 15 or 16 years old spend twice as much time online as 11-year-olds.[114]

The majority of children participating in the focus groups (aged 11 to 16 years old) reported **spending three to more than five hours daily online**. Around 29% of participants declared spending more than five hours per day online.

Graph 6. Focus group answers on time spent online per day



Source: Focus groups in 10 countries across Asia, Europe and Latin America

Studies suggest that there are ages where children are more vulnerable to the negative effects of social media use, when it may cause more harm. Another source found that the percentage of children that encounter a negative experience online increases with age.[115] The onset of puberty, at 14 or 15 for boys and 11 to 13 for girls, is a period of particular vulnerability.[116] This is the exact window when most children are victimised. Research shows that **the majority of victims of commercial child sexual exploitation are between 13 to 15 years old**.[117] Changes taking place during early adolescence in terms of physical, social, and cognitive development make children around puberty more vulnerable in the situations of sexual encounters online compared to their older counterparts or adults.[118]

[114] Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S. and Hasebrink, U. (2020) *EU Kids Online 2020: Survey results from 19 countries*.

[115] ibid

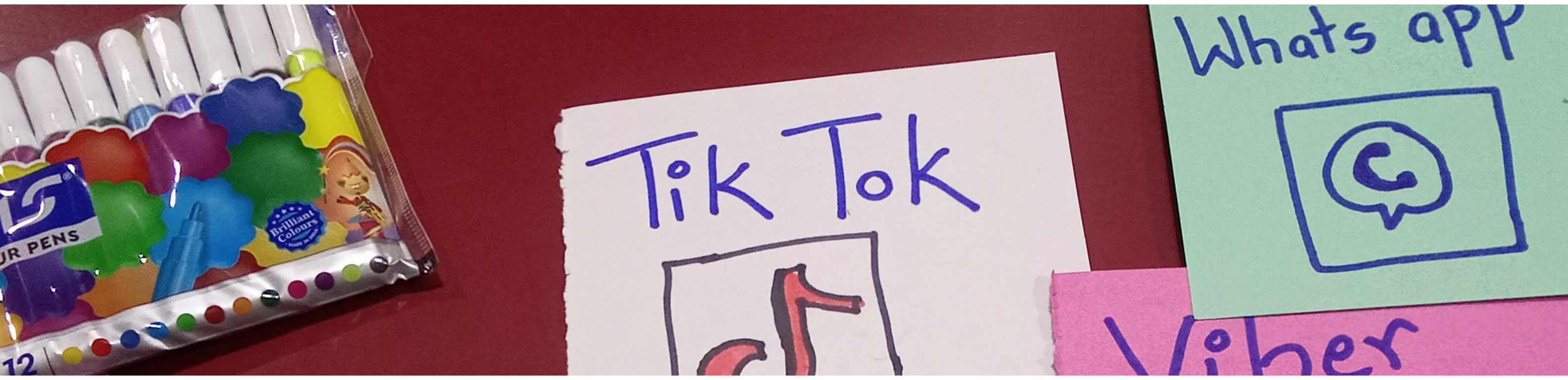
[116] Orben, A., Przybylski, A.K., Blakemore, S.J. and Kievit, R. A. (2022) 'Windows of developmental sensitivity to social media', *Nature Communications*, 13, 1649. https://doi.org/10.1038/s41467-022-29296-3.

[117] Benavente, B., Ballester Brage, L., Pich Solé, J. and Pereda Beltrán, N. (2021) 'Risk Factors for Commercial Sexual Exploitation of Children and Adolescents: Results of an International Delphi Panel', *Psicothema*, 33(3), pp.449-455. Available at: http://www.psicothema.com/pdf/4691.pdf (Accessed 18 April 2022).

[118] Burén, J. and Lunde, C. (2018) 'Sexting Among Adolescents: A Nuanced and Gendered Online Challenge for Young People'.

Knowledge about the developmental stage of adolescence can help improve the safety of designs.[119] During teenage years and adolescence, the child's focus shifts from being a part of a family to becoming more autonomous.[120] The adolescent craves more privacy,[121] will slowly take more distance from their parents and want to focus on peers.[122] Children will actively search for more connections, because they are seeking validation and acceptance.[123] These needs increase as teenagers get older.[124] For their online activities, this means they will prefer to use their own mobile phone over shared computers for instance. Valuing their mobile phones so much, this also means that teenagers are almost always online on a device that parents have little control over.[125]

Biologically, the body changes during adolescence, which influences children's **body image and self-esteem**. Having a low self-esteem and craving the approval of others about their changing body can lead to risky sexual behaviour.[126] Additionally, more **hormones** are released by the body during adolescence, including sex hormones. This results in the development of sexual interest.[127] The internet will also become a place where adolescents can orient themselves sexually. While younger children might come across explicit online content by accident, older children will increasingly view this kind of content more intentionally as they develop their own sexual identity.[128] While this could be a healthy form of exploring, there are some risks attached. Exploring sexuality online might also mean that children happen upon CSAM on the internet. Being exposed to this type of content is associated with having negative effects on their wellbeing. Additionally, some people that have watched pornography from a young age might increasingly explore more extreme material – between 4% and 17% of young people are estimated to view violent or illegal pornography, often as a result of frequent viewing of pornography.[129] Becoming desensitised to extreme content such as CSAM could also lower the threshold for continuing to watch CSAM in adulthood. In a 2021 survey of people using CSAM, it was found that 40% viewed CSAM when they were below the age of 13. Of users surveyed, 70% saw this type of content before reaching adulthood.[130]



Source: Focus group in Nepal

[119] Badillo-Urquiola, K., Chouhan, C., Chancellor, S, De Choudhary, M. and Wisniewski, P. (2019) 'Beyond Parental Control: Designing Adolescent Online Safety Apps Using Value Sensitive Design', *Journal of Adolescent Research*, 35(1), pp.147-175. https://doi.org/10.1177/0743558419884692.

[120] ibid

[121] Burén, J. (2020) *Sexting among adolescents: A gendered online phenomenon, related to individual and social determinants*. PhD thesis. University of Gothenburg.

[122] Laghi, F. and Schneider, B. (2013) 'Knowing when not to use the Internet: Shyness and adolescents' on-line and off-line interactions with friends'. *Computers in Human Behavior,* 29, pp.51-57. https://doi.org/10.1016/j.chb.2012.07.015.

[123] Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[124] Wisniewski, P., Ghosh, A. K., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Parental Control vs. Teen Self-Regulation: Is there a middle ground for mobile online safety?', *2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. Portland, United States. 25 February – 1 March. ACM, New York, United States. https://doi.org/10.1145/2998181.2998352.

[125] ibid

[126] Burén, J. (2020). *Sexting among adolescents*.

[127] ibid

[128] Nash, V., Adler, J. R., Horvath, M. A. H., Livingstone, S., Marston, C., Owen, G. and Wright, J. (2015) *Identifying the routes by which children view pornography online: implications for future policy-makers seeking to limit viewing.*

[129] NSPCC (2021) *Statistics briefing: harmful sexual behaviour*. Available at: https://learning.nspcc.org.uk/media/1661/statistics-briefing-harmful-sexual-behaviour.pdf. (Accessed 25 April 2022).

[130] Insoll, T., Ovaska, A. and Vaaranen-Valkonen, N. (2021) *CSAM Users in the Dark Web: Protecting Children Through Prevention*. Available at: https://drive.google.com/file/d/1EUBsU0A8XYw8QNUg3JKqemIRoLO9cYPt/view (Accessed 18 April 2022).

While biological development is fast during the teenage years, **cognitive development** is slower. The prefrontal cortex, which is responsible for controlling impulses, is less developed until the mid-twenties. This has an influence on making decisions and overseeing the consequences.[131] The combination of wanting more autonomy, having more hormones in the body, and the **lack of impulse control and self-monitoring abilities** can thus result in more risk-taking.[132] Teenagers' developmental stage could therefore result in them engaging in risky situations online such as sexting. In a study of adolescents seeking support about sexting, Razi, Badillo-Urquiola and Wisniewki (2020) found that adolescents used sexting to take a relationship further. Here, the distinction between consensual and non-consensual sexting comes into play. When consensual, sexting could be seen as a natural progression of a healthy romantic relationship.[133] However, when there is no consent, sexting could place teenagers at online and offline risks, such as photos being distributed to other peers, bullying, or feeling forced into doing things they did not want to.[134]

"There were some cases I heard of in which some girls found their pictures and videos shared online, without their consent." (Focus group discussions, 15-year-old girl, Romania)

"A friend of mine sent a picture to her boyfriend wearing only a bra, and when they broke up, he published it on Facebook. She was ashamed to go out to the Street, until she talked to her mother, and she supported her." (Focus group discussions, 14 year old girl, Nicaragua)

An Economist Impact global survey found that almost one in five (18%) young people reported having a "sexually explicit image of themselves being shared by a peer without consent".[135] When explicit photos are shared and the other person is threatening to disseminate them further, this is called sexual extortion.[136] In addition, sexting makes teenagers vulnerable to offline harms, including bullying and sexual predation.[137]

Children's digital skills can also influence the risks they face online. Higher digital skills combined with risk-taking behaviour increases the likelihood of facing online risks.[138]

2.1.4 Disability increases the risks of OSEC

Children with disabilities are more likely to experience abuse, including OSEC. Factors contributing to the vulnerability of children with disabilities include poverty, neglect, extensive time online compared to peers without disabilities, isolation, gender (girls being at higher risk of abuse), failure of adequate protection measures, and lack of access to education.[139] Children with disabilities often lack access to education, in particular sex education, as they may be deemed incapable of learning or due to the fear that access to sex education could promote sexual behaviour. Because of the lack of access to sex education, children with disabilities may have more difficulties in identifying the risks of grooming or other forms of OSEC.[140]

[131] Burén, J. (2020) ***Sexting among adolescents***.

[132] Badillo-Urquiola, K., Chouhan, C., Chancellor, S, De Choudhary, M. and Wisniewski, P. (2019) 'Beyond Parental Control: Designing Adolescent Online Safety Apps Using Value Sensitive Design'; Burén, J. (2020). ***Sexting among adolescents***; Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[133] Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2020) 'Let's Talk about Sext: How Adolescents Seek Support and Advice about Their Online Sexual Experiences', ***2020 CHI Conference on Human Factors in Computing Systems***, Honolulu, United States, 25-30 April. ACM, New York, United States. **https://doi.org/10.1145/3313831.3376400**.

[134] Tariq, M. U., Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2019) 'A Review of the Gaps and Opportunities of Nudity and Skin Detection Algorithmic Research for the Purpose of Combatting Adolescent Sexting Behaviors'. ***21st International Conference on Human-Computer Interaction***. Orlando, United States, 26-31 July. Springer, Cham. **https://doi.org/10.1007/978-3-030-22636-7_6**.

[135] Economist Impact and WeProtect (2022) ***Estimates of childhood exposure to online sexual harms and their risk factors***.

[136] Patchin, J. W. and Hinduja, S. (2018) 'Sextortion Among Adolescents: Results From a National Survey of U.S. Youth', ***Sexual Abuse***, 32(1), pp.30-51. **https://doi.org/10.1177/1079063218800469**.

[137] Tariq, M. U., Razi, A., Badillo-Urquiola, K. A. and Wisniewski, P. (2019) 'A Review of the Gaps and Opportunities of Nudity and Skin Detection Algorithmic Research for the Purpose of Combating Adolescent Sexting Behaviors'.

[138] Livingstone, S. and Smith, P. K. (2014) 'Annual research review: Harms experienced by child users of online and mobile technologies: the nature, prevalence and management of sexual and aggressive risks in the digital age'.

[139] WeProtect Global Alliance (2021) The sexual exploitation and abuse of deaf and disabled children online. Available at: **https://www.weprotect.org/wp-content/uploads/Intelligence-briefing-2021-The-sexual-exploitation-and-abuse-of-disabled-children.pdf** (Accessed 7 May 2022).

[140] ibid

Loneliness and isolation may also influence children with disabilities to accept sexual requests in return for attention and social acceptance. The stigma around disabilities and sexual needs makes it harder for adolescents with disabilities to safely explore their sexuality online, and contributes to their vulnerability.[141] Children with disabilities, especially intellectual impairments, may have lower digital skills but they are more susceptible to online grooming. According to Global Threat Assessment 2021, a disproportionate number of respondents who identified as disabled experienced online sexual harm by an adult they knew.[142]

2.1.5 Children are a diverse group and intersecting factors contribute to risks

The literature reports other psychological and socio-economic risk factors for OSEC. These include low self-esteem, mental health issues,[143] disruptive or dysfunctional family dynamics, struggles with social interactions,[144] and impulsivity.[145] Factors increasing the risks of child sexual exploitation according to a systematic review of existing studies include emotional dysregulation, suicidality, growing up in a single-parent household, history of homelessness, poverty and criminality within the household, childhood trauma, post-traumatic stress disorder or being exposed to CSAM.[146] Furthermore, racial or ethnic minorities are also more likely to experience sexual harm online.[147]

Children staying in foster care and runaways are also especially vulnerable to experiencing exploitation.[148] At the same time, trauma of experiencing sexual abuse an at early age may trigger development of atypical regulatory strategies such as running away.[149] Lack of adequate documentation is also considered a risk factor, making children in migration (especially undocumented minors) vulnerable to exploitation.[150]

The COVID-19 pandemic has contributed to an increased vulnerability of children in a variety of ways. For example, children were confined in homes, which are not always safe spaces (research shows that most abuse happens in the closest circle), while school closures limited access to support and increased children's unsupervised screen time and the activity of young children on digital platforms.[151] Additionally, the pandemic has led to more feelings of loneliness. This resulted in children seeking connections online with friends, but also with strangers.[152]

[141] ibid

[142] WeProtect Global Alliance (2021) ***Global Threat Assessment 2021 – Working together to end the sexual abuse of children online***, p.39. Available at: **https://www.weprotect.org/global-threat-assessment-21/#report** (Accessed 18 April 2022).

[143] Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[144] Schoeps, K., Peris Hernández, M. P., Garaigordobil, M., & Montoya-Castilla, I. (2020). Risk factors for being a victim of online grooming in adolescents. ***Psicothema,*** 31 (4); Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[145] Schoeps, K., Peris Hernández, M. P., Garaigordobil, M. and Montoya-Castilla, I. (2020) 'Risk factors for being a victim of online grooming in adolescents'. ***Psicothema***, 32(1), pp.15-23. **https://doi.org/10.7334/psicothema2019.179**.

[146] Laird, J., Klettke, B., Hall, K., Clancy, E. and Hallford, D. (2020) 'Demographic and Psychosocial Factors Associated With Child Sexual Exploitation – A Systematic Review and Meta-analysis'.

[147] Economist Impact and WeProtect Global Alliance (2022) ***Estimates of childhood exposure to online sexual harms and their risk factors***; Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[148] Benavente, B., Ballester Brage, L., Pich Solé, J. and Pereda Beltrán, N. (2021) 'Risk Factors for Commercial Sexual Exploitation of Children and Adolescents: Results of an International Delphi Panel'; Badillo-Urquiola, K., Page, X. and Wisniewski, P. (2019) 'Risk vs. Restriction: The Tension between Providing a Sense of Normalcy and Keeping Foster Teens Safe Online', ***Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems***. ACM, New York, United States, Paper 267, pp.1-14. **https://doi.org/10.1145/3290605.3300497**.

[149] ibid

[150] ibid

[151] UNICEF (2021) ***Ending online child sexual exploitation and abuse: Lessons learned and promising practices in low- and middle-income countries***.

[152] WeProtect Global Alliance (2021) ***Global Threat Assessment 2021 – Working together to end the sexual abuse of children online***.

2.1.6 Children's rights in the digital environment

Undeniably, children's lives exist in both online and offline environments. Children have grown accustomed to learning, playing, and connecting with others online, which has been further exacerbated by the COVID-19 pandemic. Clearly defining the rights of the child in the digital environment is more relevant than ever. In 2021, the UN Committee on the Rights of the Child (CRC) adopted its General comment No. 25 on the rights of the child in the digital environment. The Comment provides four general principles in line with the United Nations Convention on the Rights of the Child (UNCRC) to ensure the rights of children in the digital environment:

1. Non-discrimination (Article 2)
2. Best interests of the child (Article 3)
3. Right to life, survival, and development (Article 6)
4. Respect for the views of the child (Article 12)[153]

The rights of the child in the online environment should be protected and fulfilled in the same way they would be in the offline world. The digital environment offers opportunities for children to exercise their civil rights and freedoms such as their right to privacy and freedom of expression, and their right to protection from all kinds of online violence.

> Children consulted by the UN CRC and 5Rights Foundation expressed, in simple terms, how States can protect and fulfil their rights in the digital environment: "***make rules*** *that support children's rights online, ensure that everyone understands these rules,* ***listen*** *to children when they have a problem,* ***stop*** *business* ***putting profit*** *above children's rights, make sure that there are* ***consequences*** *for organisations or people who break the rules, and lastly, make sure children* ***know*** *and understand what action has been taken"*.[154]

It is important to protect children and their rights as digital citizens and content creators. Under General Comment No. 25, States are urged to take measures to ensure that businesses prevent their networks or online services from being used in ways that cause or contribute to violations or abuses of children's rights.[155] Within the principle of non-discrimination, General comment No. 25 calls for proactive **prevention measures against child sexual exploitation**, with particular attention to prevent the access to and dissemination of *"gender-stereotyped, discriminatory, racist, violent, pornographic and exploitative information"*. Safety and protective measures should be implemented taking into account gender and intersecting grounds of discrimination and in accordance with children's evolving capacities.

The General comment No. 25 also asks that relevant laws and policies ensure children are protected against economic, sexual, and other forms of exploitation, including *"emerging risks of all forms of violence in the digital environment"*. States must also ensure that appropriate enforcement mechanisms are in place which are easily accessible for children, parents, and caregivers.

In the context of child safety by design, businesses should be required to implement regulatory frameworks, industry codes, and terms of services that adhere to the **highest standards of ethics, privacy and safety** in relation to the design, engineering, development, operation, distribution and marketing of their product and services. This may also include providing child-friendly and age-appropriate explanations of terms of services of their platform. States should conduct a child rights impact assessment (CRIA)[156] when developing laws and policies and when making any decisions concerning children. In 2014, only five EU Member States were conducting such assessments: Austria, Belgium, France, Italy and Sweden.[157] The CRIA should, in addition, be published and must be available for all stakeholders.

[153] UN CRC (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment*. Available at: **https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation** (Accessed: 27 April 2022).

[154] 5Rights Foundation (n.d.) *In our own words – children's rights in the digital world*. Available at **https://5rightsfoundation.com/In_Our_Own_Words_Young_Peoples_Version_Online.pdf** (Accessed 22 April 2022).

[155] UN CRC (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment*, paragraph 36.

[156] *"Child rights impact assessment is a tool predicting the impact of any proposed law, policy or budgetary allocation, which affects children and the enjoyment of their rights."* European Union Agency for Fundamental Rights (n.d.) Child rights impact assessment. Available at: **https://fra.europa.eu/en/content/child-rights-impact-assessment#:~:text=Child%20rights%20impact%20assessment%20is,development%20of%20policies%20and%20laws** (Accessed 25 April 2022).

[157] ibid

2.1.7 The mechanisms through which risk occurs

INTERACTIONS BETWEEN CHILDREN AND UNKNOWN ADULTS

Adult predators can take advantage of risky behaviour together with additional vulnerabilities that adolescents still have.[158] **Online grooming** can be a long process that requires a lot of effort from the predator. Therefore, these predators are mostly calculating individuals that are not likely to be impulsive.[159] The course of conversation for grooming can, however, differ per perpetrator. Some use a longer-term approach, by building rapport with the child. Using flattery, they will slowly introduce sexual topics while making sure the child is not scared away.[160] Razi et al. (2021) call this the Luring Communication Theory.[161] If the child is reluctant, the predator will use tactics such as pretending to be disappointed that the child does not want to communicate with them in this way. Other predators use a more direct approach where they talk about sex in the early stages. They also tend to be more aggressive in their tactics such as using threats and insults.[162]

The most commonly used model for grooming as proposed by O'Connell (2003) suggests that the process of grooming consists of five steps:

1. friendship forming,
2. relationship forming,
3. risk assessment,
4. exclusivity, and
5. sexual stages.[163]

In one of the key articles about sex offenders, it was found that an **assessment of the vulnerability** leads to the decision of who to victimise. Sexual predators will assess which children are likely to be vulnerable and who are most likely to follow through with contact with them.[165] Section 2.2 discusses these various risk factors for being groomed. Together with factors such as availability and distance, the victim will be chosen.[166] Other scholars also found that this risk assessment is very important for predators and might even be used in the first moments of contact between a predator and a child. In one research of conversations between predators and children, in 30% of the conversations, offenders already asked questions that allowed them to assess the risk, such as the schedule of the parents, for instance. They therefore propose that this model should not be seen as a linear model, but rather as different steps that predators can take at different occasions or simultaneously.[167] One of the children in the focus groups also experienced this:

> *"Men or women [...] tell me to give them the address of my home. I do not know them and when I give it to them, they can come in!"* (Focus group discussions, boy, Nicaragua)

It has become increasingly acceptable for young people to have contact with people they do not know on the internet, making it easier for children to seek unknown contacts online without parents being worried.[168] Around 42% of children in the focus groups disagreed that it is safe to chat with someone they do not know provided they are careful. A slightly lower percentage (34%) of the participating children found this to be safe (see Graph 7). The concept of being careful is, however, broad. Multiple children mentioned that as long as they would not share personal details, they felt they should be safe.

[158] Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[159] ibid

[160] Black, P. J., Wollis, M., Woodworth, M. and Hancock, J. T. (2015) 'A linguistic analysis of grooming strategies of online child sex offenders: Implications for our understanding of predatory sexual behavior in an increasingly computer-mediated world'. ***Child Abuse & Neglec***t, 44, pp.140-149. **https://doi.org/10.1016/j.chiabu.2014.12.004**.

[161] Razi, A., Kim, S., Alsoubai, A., Stringhini, H., Solorio, T., De Choudhury, M. and Wisniewski, P. (2021) 'A Human-Centered Systematic Literature Review of the Computational Approaches for Online Sexual Risk Detection', ***Proceedings of the ACM on Human-Computer Interaction***, 5(CSCW2), 465, pp.1-38. **https://doi.org/10.1145/3479609**.

[162] Black, P. J., Wollis, M., Woodworth, M. and Hancock, J. T. (2015) 'A linguistic analysis of grooming strategies of online child sex offenders: Implications for our understanding of predatory sexual behavior in an increasingly computer-mediated world'.

[163] ibid

[164] Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[165] ibid

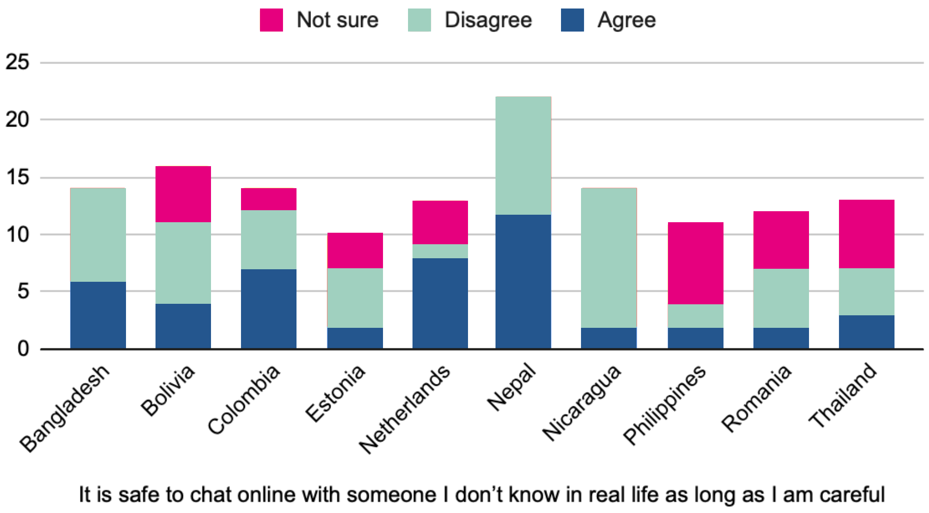
[166] ibid

[167] Black, P. J., Wollis, M., Woodworth, M. and Hancock, J. T. (2015) 'A linguistic analysis of grooming strategies of online child sex offenders: Implications for our understanding of predatory sexual behavior in an increasingly computer-mediated world'.

[168] ibid

"I don't share private info, so what's the harm?" (Focus group discussions, girl, Romania)

Graph 7. Focus group answers on safety of chatting online



Source: Focus groups in 10 countries across Asia, Europe and Latin America

Children being in contact with a predator online does also not necessarily lead to victimisation. Research suggests that children are often very much able to respond in an appropriate manner or to leave a conversation when they are uncomfortable.[169] This does not mean that online contact is without risk for children. Sometimes, grooming behaviour is not recognised as such by children. Children or adolescents engaging in sexting might think that sending pictures to strangers is a voluntary act that they think nothing of, while it might be due to manipulation by adults.[170] This also came back in the focus group discussions. As a girl from Colombia mentioned:

"There is no danger because it is by chat, they can't get to you in real life. As long as I don't give out my personal information." (Focus group discussions, girl, Colombia)

Children in multiple focus groups expressed wanting to give strangers the benefit of the doubt. They indicated that they are aware of potentially bad intentions, but that they cannot assume this is the case straights away. In the Netherlands, children were not sure what to think, because they also wanted to have faith in the fact that there are many people on the internet that do not have bad intentions. Similarly in Nepal, children expressed that it is difficult to assess the trustworthiness of a person met online as you cannot know their intentions and behaviour. In Thailand, one of the children worded this as follows:

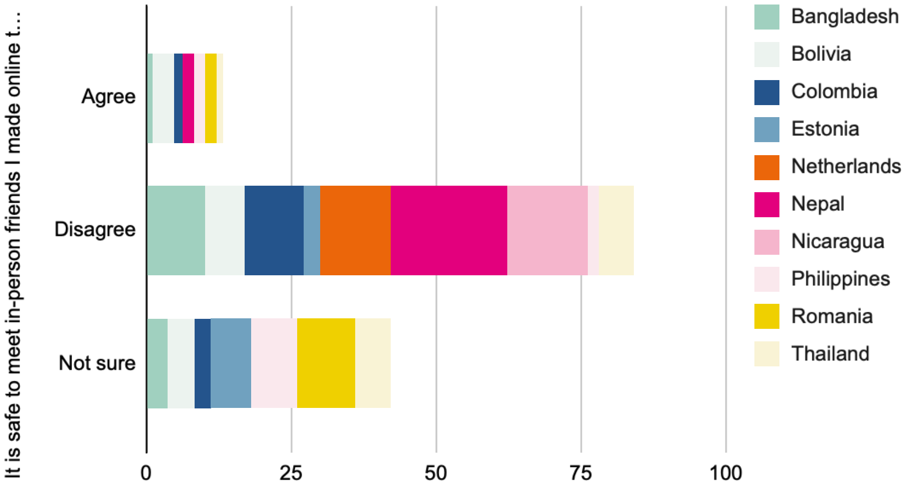
"I cannot really judge from the way the person talked to me online. I cannot judge whether the person is sincere or not. Some may want to lure me to meet them in person in order to do something bad to me. Nonetheless, as said earlier, some may be good persons." (Focus group discussions, Thailand)

When asked if they believe it is safe to meet in person with friends made online that they have never met before, most of the children disagreed with the statement.

[169] Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming'.

[170] De Santisteban, P. and Gámez-Guadix, M. (2017) 'Prevalence and Risk Factors Among Minors for Online Sexual Solicitations and Interactions With Adults'. *The Journal of Sex Research*, 55(7), pp.939-950. https://doi.org/10.1080/00224499.2017.1386763.

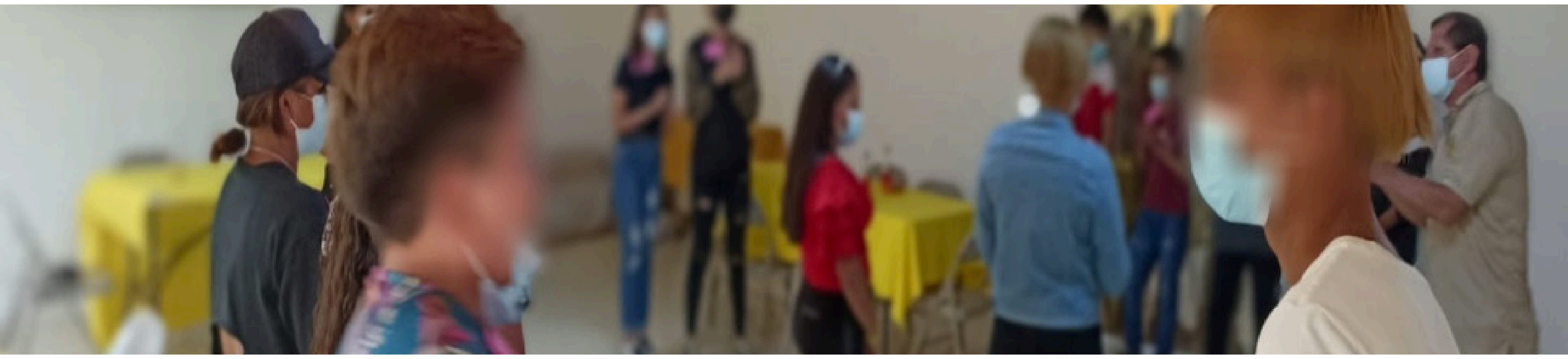
Graph 8. Focus group answers on safety of meeting online friends in person



Source: Focus groups in 10 countries across Asia, Europe and Latin America

Important to note is that this manipulation, but also grooming and abuse in general, is not limited to unknown adults. The fear of online risks is often for dangerous strangers.[171] Research suggests, however, that young people are most often victimised by family, friends or acquaintances. A reason for this could be that adolescents find it easier to shut down unwanted conversations with strangers than with people they know.[172] Blocking an unknown person, for instance, could work with strangers. With known people this is not a realistic solution and often poses a dilemma in how to handle the situation.[173]

Graph 9 and 10 show that across the focus group countries, children generally felt similarly about saying no to either friends or adults; most children agree that it is easy saying no, except for children in Bangladesh, where the majority finds it difficult to say no to adults and they are unsure whether they find it easy to say no to friends. In Bolivia and Romania, children are mostly unsure if they find it easy to say no to adults, but do have an opinion when it comes to friends. In Estonia, the Philippines, and Romania, more children agreed that it is easier to say no to friends compared to other countries. The children in the focus groups thus had different opinions about this.



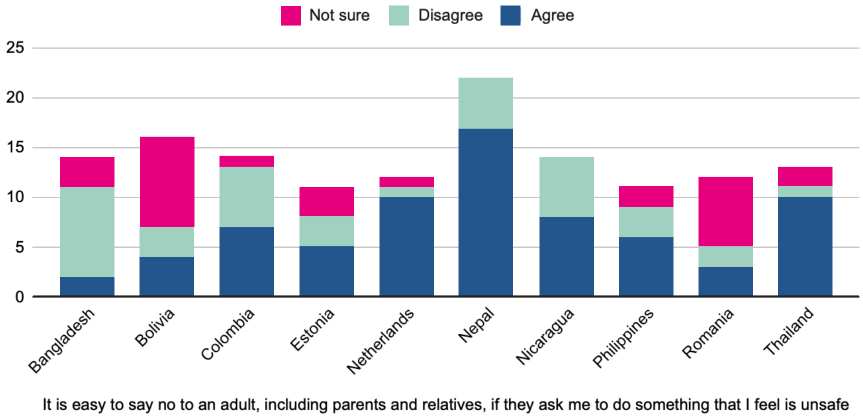
Source: Focus group in Nicaragua

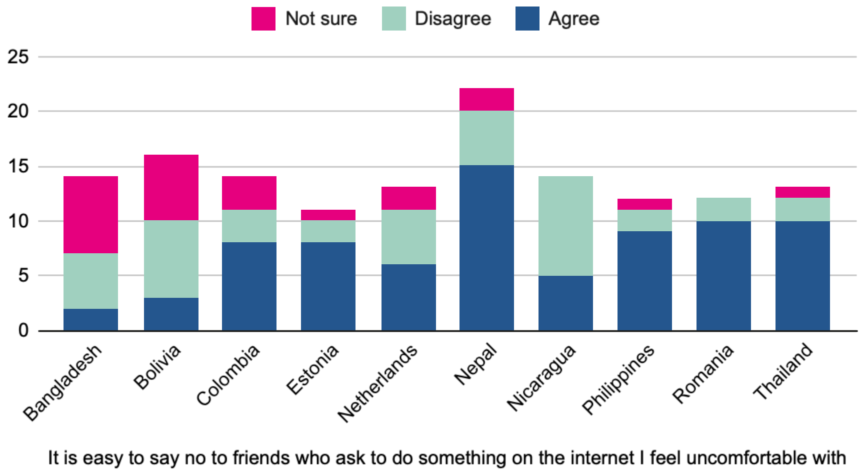
[171] Hartikainen, H., Razi, A. and Wisniewki, P. (2021) 'Safe Sexting: The Advice and Support Adolescents Receive from Peers Regarding Online Sexual Risks'. ***Proceedings of the ACM on Human-Computer Interaction***, 5(CSCW1), 42, pp.1-31 https://doi.org/10.1145/3449116; Finkelhor, D., Jones, L. and Mitchell, K. (2021) 'Teaching Privacy: A flawed strategy for children's online safety'. ***Child Abuse & Neglect***, 117. https://doi.org/10.1016/j.chiabu.2021.105064.

[172] Hartikainen, H., Razi, A. and Wisniewski, P. (2021) 'Safe Sexting: The Advice and Support Adolescents Receive from Peers Regarding Online Sexual Risks'.

[173] Sultana, S., Deb, M., Bhattacharjee, A., Hasan, S., Raihanul Alam, S. M., Chakraborty, T., Roy, P., Fairuz Ahmed, S., Moitra, A., Ashraful Amin, M., Najmul Islam, A. K. M. and Ishtiaque Ahmed, S. (2021) ''Unmochon': A Tool to Combat Online Sexual Harassment over Facebook Messenger'. ***CHI Conference on Human Factors in Computing Systems*** (CHI '21), Yokohama, Japan, 8-13 May, pp.1-18. https://doi.org/10.1145/3411764.3445154.

Graph 9 and 10. Focus group answers on saying no to adults or friends





Source: Focus groups in 10 countries across Asia, Europe and Latin America

INTERACTIONS BETWEEN CHILDREN

Children may themselves commit OSEC against other children, including through viewing CSAM, self-generating CSAM, grooming or sexual harassment, or other forms of OSEC. Studies report that a substantial percentage of child sexual harm and abuse is committed by people under 18 years old, including online.[174] Harmful sexual behaviour is most commonly displayed by adolescent boys, although girls and young children can also display such behaviour.[175] Estimates vary according to the studies, with figures suggesting that young people account for 3% to 15% of access to indecent images of children.[176] Based on US national crime data, a study found that youth offenders constituted about 35% of all sex crime cases against children and about 50% of cases involving a child under the age of 12.[177] Similar studies found various ranges from one third to 40%, 50% and two thirds of child offenders in sexual offences against other children.[178] Official case data in the UK

[174] Letourneau, E. J., Schaeffer, C. M., Bradshaw, C. P. and Feder, K. A. (2017) 'Preventing the Onset of Child Sexual Abuse by Targeting Young Adolescents With Universal Prevention Programming', *Child Maltreatment*, 22(2), pp.100–111. https://doi.org/10.1177/1077559517692439.

[175] NSPCC (2021) *Statistics briefing: harmful sexual behaviour.*

[176] Belton, E. and Hollis, V. (2016) *A Review of the Research on Children and Young People who Display Harmful Sexual Behaviour Online.* Available at: https://learning.nspcc.org.uk/media/1198/review-children-young-people-harmful-sexual-behaviour-online.pdf (Accessed 25 April 2022).

[177] Letourneau, E. J., Schaeffer, C. M., Bradshaw, C. P. and Feder, K. A. (2017) 'Preventing the Onset of Child Sexual Abuse by Targeting Young Adolescents With Universal Prevention Programming'.

[178] Letourneau, E. J., Schaeffer, C. M., Bradshaw, C. P. and Feder, K. A. (2017) 'Preventing the Onset of Child Sexual Abuse by Targeting Young Adolescents With Universal Prevention Programming'; Sneddon, H., Gojkovic Grimshaw, D., Livingstone, N. and Macdonald, G. (2020) 'Cognitive-behavioural therapy (CBT) interventions for young people aged 10 to 18 with harmful sexual behaviour', *Cochrane Database of Systematic Reviews*, 6, CD009829. https://doi.org/10.1002/14651858.CD009829.pub2.

show that "*about one in 1000 young people aged 12 to 17 years old is identified as displaying [harmful sexual behaviour]*", while in Germany, criminal data indicate young people aged 14 to 20 years are "over-represented in the category of 'sexual offences'.[179] Age and gender affect the likelihood of sexually harmful behaviour by children, with early adolescence being identified as a 'peak time' for displaying such behaviours. Boys are much more likely to exhibit such behaviours; some studies indicate that between 92% and 97% of young people displaying sexually harmful behaviours are boys while 3% to 8% are girls.[180] The NSPCC has found some cross-over between online and offline harmful sexual behaviour as well as between harmful sexual behaviour and child sexual exploitation.[181]

One of the difficulties in having precise estimates lies in **under-reporting** of offences. A UK prevalence study reported that "*83% of young people aged 11 to 17 years old who had been sexually assaulted by a peer had not told anyone about the assault*". In comparison, the non-disclosure rate of adult victims is 34%. Under-reporting may also be more likely in communities and cultures where sexuality is taboo or shameful, with a highly victim-blaming attitude.[182] In the focus group in Nepal, this was discussed. The children mentioned that if anything uncomfortable happened on the internet, they did not want to share it with anyone. They also indicated that they were unaware of how reporting works.

Internet Watch Foundation (IWF) has disclosed that one third of CSAM online reported in the UK is self generated by children themselves.[183] The COVID-19 pandemic led to a sharp increase of **self-generated CSAM** content from 2019 to 2020.[184] This could mean explicit content created by children using their mobile phones or webcams.[185] ECPAT's Terminology Guidelines warn about the use of this terminology, as it could imply that the child is at fault or responsible for the content.[186] However, children could be forced by others to produce or share their content or (child) offenders could share the self-produced content without the child knowing.[187] This can result in the viewing and recirculation of the content, re-victimising the child as long as the content is online.[188]

While **sexting** can be a healthy exploration of sexuality, it can also lead to harmful behaviours such as producing, viewing and disseminating CSAM or sexual exploitation material. Sexting is linked to sexually risky behaviours. In fact, *"[b]oth experimental and abusive circumstances have been identified in which young people sext"*.[189]

Studies indicate that it is **common for girls to have received unsolicited sexts or unwanted requests** for sexts from a partner, peer or known person.[190] A recent study in Spain shows that 48% of girls aged 14 to 20 years old received sexually explicit pictures from boys and around 44% requested such material from boys, while about 17% of boys the same age recognised requesting such material online and 10% reported having sent pictures to girls.[191] When requests for sexting and unsolicited sexts come from peers, partners, or friends of

[179] Sneddon, H., Gojkovic Grimshaw, D., Livingstone, N. and Macdonald, G. (2020) 'Cognitive-behavioural therapy (CBT) interventions for young people aged 10 to 18 with harmful sexual behaviour', *Cochrane Database of Systematic Reviews*, 6, CD009829. https://doi.org/10.1002/14651858.CD009829.pub2.

[180] ibid

[181] NSPCC (2021) ***Statistics briefing: harmful sexual behaviour***.

[182] Sneddon, H., Gojkovic Grimshaw, D., Livingstone, N. and Macdonald, G. (2020) 'Cognitive-behavioural therapy (CBT) interventions for young people aged 10 to 18 with harmful sexual behaviour'.

[183] Internet Watch Foundation (2020) ***Face the facts – The Annual Report 2020***. Available at: https://annualreport2020.iwf.org.uk/ (Accessed 18 April 2022).

[184] Internet Watch Foundation (2020) ***Trend: 'Self-generated' content***. Available at: https://annualreport2020.iwf.org.uk/trends/international/selfgenerated (Accessed 25 April 2022).

[185] ibid

[186] ECPAT International (2016) ***Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse (commonly known as the 'Luxembourg Guidelines')***, p.17. Available at: https://ecpat.org/wp-content/uploads/2021/05/Terminology-guidelines-396922-EN-1.pdf (Accessed 18 April 2022).

[187] Internet Watch Foundation (2020) ***Trend: 'Self-generated' content***.

[188] Insoll, T., Ovaska, A. and Vaaranen-Valkonen, N. (2021) ***CSAM Users in the Dark Web: Protecting Children Through Prevention***.

[189] Belton, E. and Hollis, V. (2016) ***A Review of the Research on Children and Young People who Display Harmful Sexual Behaviour Online***.

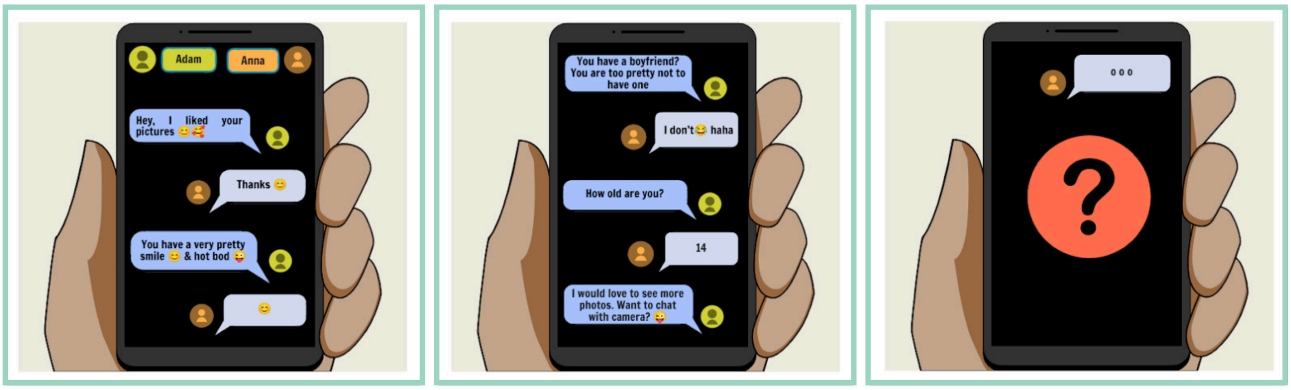
[190] Mishna, F., Milne, E., Cook, C., Slane, A. and Ringrose, J. (2021) 'Unsolicited Sexts and Unwanted Requests for Sexts: Reflecting on the Online Sexual Harassment of Youth'.

[191] Díaz-Aguado, M. J. and Martínez-Arias, R. (2022) 'Types of Male Adolescent Violence Against Women in Three Contexts: Dating Violence Offline, Dating Violence Online, and Sexual Harassment Online Outside a Relationship'.

a mutual contact, young people report finding it difficult to decline such requests.[192] Research indicates a negative impact of unsolicited sexts, leading to higher depression and anxiety, and lower self-esteem.[193]

In the focus group discussions, the children were shown a chat conversation between 'Adam' and 'Anna' (see Figure 1), where Adam compliments Anna on her pictures and requests a video chat. The children responded differently to this chat. Usually a small part of the group expressed that this was a normal way of starting a conversation. This might indicate that these children come across these requests more often. In general, a little over half of the children in the focus groups saw some red flags in this conversation. In responding to the question of what Anna should do next, most children suggested that they stop chatting, block or report the person, or go to a trusted adult or the police.

Figure 1. Example of online chat discussed in focus group


Adam
Anna
Hey, I liked your pictures 😊🥰
Thanks 😊
You have a very pretty smile 😊 & hot bod 😜
😊
You have a boyfriend? You are too pretty not to have one
I don't😂 haha
How old are you?
14
I would love to see more photos. Want to chat with camera? 😜
o o o
?

Create your own at Storyboard That

Source: Focus groups in 10 countries across Asia, Europe and Latin America

2.2 SAFETY MEASURES CURRENTLY IN PLACE

This section discusses the safety measures that are used by the popular social media apps Facebook/Instagram, Snapchat, TikTok, and YouTube. It aims to provide an overview of what is currently in place before discussing in detail the solutions that are found to be effective in keeping children safe against OSEC (Section 3).

Safety measures currently in place, and reviewed for this report, include the following:

1. Community guidelines
2. Minimum age to use the platform
3. Age verification
4. Automatic settings for children
5. Classifiers
6. Deleting and reporting of material
7. Deterrence and warning messaging

These different categories will first be explained in this order in the section 'Existing safety measures used by platforms', along with both their benefits and challenges. Subsequently, the section 'Safety measures used by popular social media apps' will give an overview of which measures the popular social media platforms use.

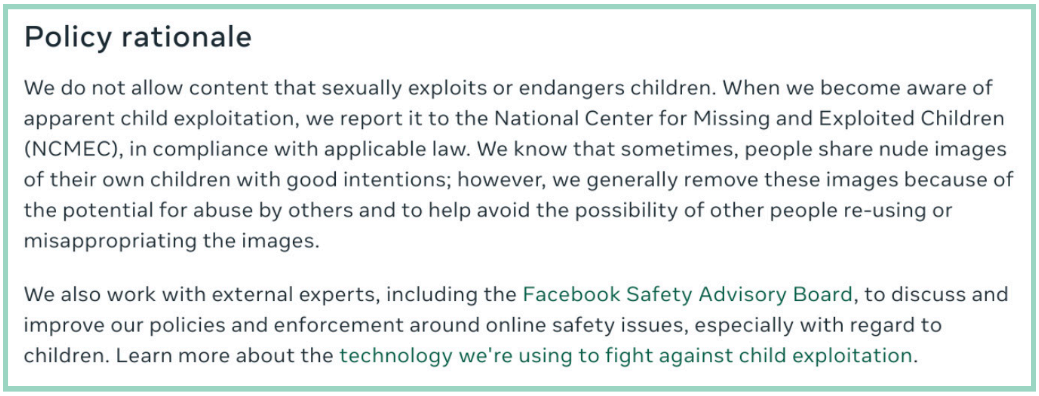
[192] Mishna, F., Milne, E., Cook, C., Slane, A. and Ringrose, J. (2021) 'Unsolicited Sexts and Unwanted Requests for Sexts: Reflecting on the Online Sexual Harassment of Youth'.
[193] ibid

2.2.1 Existing safety measures used by platforms

COMMUNITY GUIDELINES

All social media platforms have documents such as Terms and Conditions, **Terms of Use** Agreements and Codes of Conduct. These documents lay out the platforms 'dos and don'ts'. They documents also equip the social media platforms with the mandate to remove inappropriate content. Figure 2 shows the example of how Meta (Facebook, Instagram, and WhatsApp) deals with explicit content of children, even stating that they will remove well-intended photos of children to prevent wrongful use of these images.[194]

Figure 2. Meta policy on sexual exploitation of children[195]



Policy rationale

We do not allow content that sexually exploits or endangers children. When we become aware of apparent child exploitation, we report it to the National Center for Missing and Exploited Children (NCMEC), in compliance with applicable law. We know that sometimes, people share nude images of their own children with good intentions; however, we generally remove these images because of the potential for abuse by others and to help avoid the possibility of other people re-using or misappropriating the images.

We also work with external experts, including the Facebook Safety Advisory Board, to discuss and improve our policies and enforcement around online safety issues, especially with regard to children. Learn more about the technology we're using to fight against child exploitation.

Source: Facebook Community Standards – Child Sexual Exploitation, Abuse and Nudity

Meta's Community Standards are comprehensive and cover a wide range of OSEC. They include grooming and CSAM, as well as sexualised material showing children in a sexualised context, including digitally-created depictions of nude children, unless for health or educational purposes. Meta Community Standards also cover 'inappropriate interactions with children' such as:

Arranging or planning real-world sexual encounters with children
Purposefully exposing children to sexually explicit language or sexual material
Engaging in implicitly sexual conversations in private messages with children
Obtaining or requesting sexual material from children in private messages
Exploitative intimate imagery and sextortion

While the Community Standards appear comprehensive, this approach puts the responsibility on the users, by emphasising self-regulation with regard to inappropriate behaviour. This is, however, **not enough to keep young users safe**.[196] The Community Standards are broad enough to encompass many types of content, yet Meta does not appear equipped to adequately police its own platforms. Efforts to have groups targeting children under 13 years old (which are not allowed on Meta's platforms) removed have been met with inaction from Meta. During her efforts to have such groups removed, one research professor was actually suggested new child sexualisation groups as "Groups You May Like" by Meta's algorithm.[197]

In addition, the **terms and conditions are usually not child-friendly**, as the documents can be long and mainly textual. The text itself is technical, which is not very accessible to children. Children do not take the time to read the text fully and just click on agree. As a result, they do not know what they have agreed to.[198] Teenagers

[194] Meta (2022) *Facebook Community Standards – Child Sexual Exploitation, Abuse and Nudity*. Available at: https://transparency.fb.com/en-gb/policies/community-standards/child-sexual-exploitation-abuse-nudity/ (Accessed 25 April 2022).

[195] ibid

[196] Equality Now, TrustLaw and Thomson Reuters Foundation (2021) *Ending Online Sexual Exploitation and Abuse of Women and Girls: A Call for International Standards.* Available at: https://equalitynow.storage.googleapis.com/wp-content/uploads/2021/11/13160619/Ending-OSEA-Report.pdf (Accessed 25 April 2022).

[197] Putman, L. (2022) 'Facebook Has a Child Predation Problem', *Wired*, 13 March. Available at: https://www.wired.com/story/facebook-has-a-child-predation-problem/ (Accessed 7 May 2022).

[198] Milkaite, I. and Lievens, E. (2020) 'Child-friendly transparency of data processing in the EU: from legal requirements to platform policies'. *Journal of Children and Media*, 14(1), pp.5-21. https://doi.org/10.1080/17482798.2019.1701055.

might only focus on the benefits of sharing their information and want to participate in certain activities, ignoring or not being aware of potential risks.

Facebook established an Oversight Board in May 2020 to which users can appeal a decision to remove content violating its policies. In the third quarter of 2021 alone, it received 339,317 appeal cases submitted by users.[199] Meta can itself submit cases to the Board for guidance. It has done so concerning content potentially violating Facebook's Adult Nudity and Sexual Activity Community Standard. The Board cannot however be compared to judicial review as it lacks the judicial independence and power. The Oversight Board itself has criticised Facebook for its lack of transparency.[200]

2.2.2 Minimum age to use service and age verification

The minimum age to sign up for social media platforms is set at 13 years in most countries. This is due to regulations, such as the US Congress' Children's Online Privacy Protection Act (COPPA) and the EU's General Data Protection Regulation (GDPR) which set the age of **13 as the minimum age a child can consent to the processing of their personal data.**[201] Much of the focus on preventing online harm centres on age verification in relation to the risks of data processing. A wide range of age verification mechanisms exist, such as filling out an online form or taking a photo of an identification document.[202]

Most social media applications resort to **self-declaration of age** which, while cheap and easily implemented cannot be considered an effective age assurance mechanism.[203] Children under the age of 13 are likely to lie about their age knowing that otherwise they would be excluded from accessing the service. A parental consent approach used by some platforms is also not necessarily efficient, as it rarely includes verification of the guardian.[204] In addition, research shows that some parents wish to have a final say on their children's access to online services and might help children to circumvent age restrictions imposed by the apps.[205]

Social media platforms also make use of age assurance methods, where they use tools to further estimate the age of their users based on the content they make and watch, the connections they have, and the language they use.[206]

The **lack of effective age verification systems** means that children are active on platforms that are not adequate to their needs. Protective measures would also be ineffective as they would not be tailored to the actual child's age. In the focus group discussions, 38% of the children said it was okay to lie about age online (37% disagreed, 25% were unsure). The children reasoned that they sometimes had to, otherwise they were not allowed on the platform.

[199] Oversight Board (2021) *Oversight Board publishes transparency report for third quarter of 2021*. Available at: https://www.oversightboard.com/news/640697330273796-oversight-board-publishes-transparency-report-for-third-quarter-of-2021/ (Accessed 7 May 2022).

[200] Satariano, A. (2021) 'Facebook's oversight board faults its policy on preferential treatment', *New York Times*, 21 October. Available at: https://www.nytimes.com/2021/10/21/business/facebook-oversight-board-members-criticism.html (Accessed 7 May 2022).

[201] Phippen, A. and Brennan, M. (2021) *Child Protection and Safeguarding Technologies Appropriate or Excessive 'Solutions' to Social Problems?* New York: Routledge.

[202] Marley, R. (2021) 'Age Verification for Social Media- Protecting the Younger Victims of Online Scams', *ShuftiPro,* 14 October. Available at: https://shuftipro.com/blog/age-verification-for-social-media-protecting-the-younger-victims-of-online-scams/ (Accessed 25 April 2022).

[203] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls*. Available at: http://eprints.lse.ac.uk/112559/ (Accessed 22 April 2022).

[204] van der Hof, S. and Ouburg, S. (2021) *Methods for Obtaining Parental Consent and Maintaining Children Rights*. Available at: https://euconsent.eu/download/methods-for-obtaining-parental-consent-and-maintaining-children-rights/ (Accessed 25 April 2022).

[205] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls.*

[206] van der Hof, S. (2021) 'Age assurance and age appropriate design: what is required?' *LSE Parenting for a Digital Future Blog*, 17 November. Available at: https://blogs.lse.ac.uk/parenting4digitalfuture/2021/11/17/age-assurance/ (Accessed 25 April 2022).

Others indicated that they do not want to be treated like younger children online, and some even said it was safer to lie about age, because they believe abusers tend to prey on younger children.

"It is not illegal to lie about your age, everyone does it, I also lie about my age to enter online games. It's ok I always do it. It's not good, but one does it to avoid things, to be able to have access to social networks, because if you say you're older you can't be on networks and be with friends. I don't like that they know my age, I say I'm 18 years old. Yes, for security and I always do it [...] or else I wouldn't have been able to have social networks. I feel safer." (Focus group discussion, child from Colombia)

"When I play games, I choose another age so other adults who play the same game cannot see I am younger, so they cannot hurt me." (Focus group discussion, 14-year-old girl, Romania)

The children also saw the dangers of lying about age, since it could trigger algorithms that would let them see unfiltered content. In the Netherlands, the children concluded that you could lie about age when you register for apps, but they were conflicted as to whether to lie about age when you chat with someone. Some thought you should not lie about age in that context while at the same time, lying about age while chatting could be done for safety purposes.

"When you are talking to someone that is way older, like 37 and you say you are 27, they will leave you alone. For your own safety you can lie about your age". (Focus group discussion, the Netherlands)

Aside from the lack of effectiveness of age verification systems, keeping children away from platforms does not necessarily make them safe. Children find ways around being blocked and will prioritise the benefits of socialising online over the risks. Age verification and age assurance should be used in combination with designs aimed at empowering children to manage their own safety (see Section 3 on the solutions).

2.2.3 Automatic settings for children

When registering a new social media account, the content is by default set to public, meaning that all users and non-users can see all the content published. Recently, some social media platforms decided to automatically set new accounts created by children under a certain age to private, i.e. the new users' content is only shared to their friends or followers. In early 2021, TikTok decided to make the accounts of all users under the age of 16 **private by default** and changed all registered accounts of users aged between 13 and 15 years old to private.[207] Similarly, Instagram announced in July 2021 that new accounts of users under the age of 16 are set to private by default.[208] Instagram's policy now mentions that *"If you are under 18 when you sign up for an Instagram account, you'll have the option to choose between a public or private account, but private will be selected by default"*.[209]

The use of **intelligent privacy settings** can help protect children's content and information by ensuring their profiles are not accessible for the wider public.

A relevant aspect to consider when discussing user privacy choices is the existence of cognitive biases and design tactics (also called dark patterns) used by platforms to manipulate users into sharing more information. Such techniques include nudging, leading language, instant gratification linked to sharing data, or paradoxically providing an overwhelming overchoice of privacy options.[210] Even if online service providers give their users choice when it comes to data privacy, research shows that such options tend to be challenging for website users to navigate. There is a need for **more transparent information on data protection**. In order for the future designs to be effective, they should be more unified, easier to navigate (e.g. interface changes including better labelling and simple formatting) and ensure that choices made by the users will actually be honoured (e.g.

[207] TikTok (2021) *Strengthening privacy and safety for youth on TikTok*, 13 January. Available at: https://newsroom.tiktok.com/en-us/strengthening-privacy-and-safety-for-youth (Accessed 7 May 2022).

[208] Instagram (2021) *Giving Young People a Safer, More Private Experience*, 27 June. Available at: https://about.instagram.com/blog/announcements/giving-young-people-a-safer-more-private-experience (Accessed 7 May 2022).

[209] Instagram (n.d.) *Privacy settings and information*. Available at: https://help.instagram.com/196883487377501. (Accessed 7 May 2002).

[210] Waldman, A. E. (2020) 'Cognitive biases, dark patterns, and the 'privacy paradox'', *Current Opinion in Psychology*, 31, pp.105-109. https://doi.org/10.1016/j.copsyc.2019.08.025.

confirmation messaging).[211] Certain techniques implemented by designers to influence users' consent and privacy choices include salient factors such as the use of bright colours, differentiating the size and positioning of options, and preselected boxes, leading users to accept the less safe, non-privacy friendly options.[212]

Given that exercising privacy choices proves to be difficult for adult users, who could be argued to have a better understanding of such options, it is particularly important to ensure user-friendly privacy designs that keep children in mind. Furthermore, according to researchers, due to the abundance and complexity of data collection pathways, the simple **notice and consent model** is no longer sufficient to adequately inform users about their data processing.[213]

Studies show that users rarely change default settings provided by the service provider and might not even be aware of them.[214] At the same time, default settings have a significant impact on users' actions and tend to nudge them away from stronger privacy-preserving options.[215] It is therefore particularly important to ensure that such settings correspond to the safety needs of children. In addition, nudging techniques (including the use of AI systems) have been shown to have negative effects on children's digital experience, making it harder for them to disengage from apps. Considering the adverse effects this has on users, efforts should be made to limit the use of such methods.[216]

2.2.4 Automated risk detection, deterrence and reporting

CLASSIFIERS BASED ON TEXT, IMAGE OR VIDEOS

Technology exists that searches for and removes explicit content of children based on text, image, video, or live streams.[217] These are called **classifiers**.[218] In research conducted by WeProtect, 84% of the companies surveyed said they used automated or partly automated processes to detect and report cases of CSAM.[219] Existing technologies include:

- PhotoDNA, developed by Microsoft and Dartmouth, which searches for imagery of sexual exploitation of children online and reports them
- Google's Content Safety API, where companies and organisations can review explicit content at a large scale
- Facebook's open-source photo- and video-matching technology, which facilitates the safety of services and hash-sharing systems[220]

[211] Habib, H., Pearman, S., Wang, J., Zou, Y., Acquisti, A., Cranor, L., Sadeh, N. and Schaub, F. (2020) '"It's a scavenger hunt": Usability of Websites' Opt-Out and Data Deletion Choices'. *CHI Conference on Human Factors in Computing Systems* (CHI '20), Honolulu, United States, 25-30 April. ACM, New York, United States. https://doi.org/10.1145/3313831.3376511.

[212] Gray, C. M., Santos, C, Bielova, N., Toth, M. and Clifford, D. (2021) 'Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective', *CHI Conference on Human Factors in Computing Systems* (CHI '21), Yokohama, Japan, 8-13 May. ACM, New York, United States. https://doi.org/10.1145/3411764.3445779.

[213] Waldman, A. E. (2020) 'Cognitive biases, dark patterns, and the 'privacy paradox''.

[214] Norwegian Consumer Council (2018) *Deceived by design: How tech companies use dark patterns to discourage us from exercising our rights to privacy*. Available at: https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf (Accessed 25 April 2022).

[215] Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., Komanduri, S., Leon, P. G., Sadeh, N., Schaub, F., Sleeper, M., Wang, Y. and Wilson, S. (2017) 'Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online', *ACM Computing Surveys*, 50(3), 44, pp.1-41. https://doi.org/10.1145/3054926.

[216] 5Rights Foundation (2018) *Disrupted Childhood*. Available at: https://5rightsfoundation.com/static/5Rights-Disrupted-Childhood.pdf (Accessed 25 April 2022).

[217] WeProtect Global Alliance and the Technology Coalition (2021) *Findings from WeProtect Global Alliance/Technology Coalition survey of technology companies – Summary of findings*. Available at: https://www.weprotect.org/survey-of-tech-companies/ (Accessed 25 April 2022).

[218] *"In data science, a classifier is a type of machine learning algorithm used to assign a class label to a data input. An example is an image recognition classifier to label an image (e.g., "car," "truck," or "person"). Classifier algorithms are trained using labeled data; in the image recognition example, for instance, the classifier receives training data that labels images. After sufficient training, the classifier then can receive unlabeled images as inputs and will output classification labels for each image." C3ai (n.d.) Glossary*. Available at: https://c3.ai/glossary/data-science/classifier/ (Accessed 25 April 2022).

[219] WeProtect Global Alliance and the Technology Coalition (2021) *Findings from WeProtect Global Alliance/Technology Coalition survey of technology companies – Summary of findings*.

[220] Tech Coalition (2020) *The Technology Coalition Announces Project Protect – A Plan to Combat Online Child Sexual Abuse* [10 June] Available at: https://www.technologycoalition.org/newsroom/the-tech-coalition-announces-project-protect (Accessed 22 February 2022).

2.2.5 Artificial intelligence

Artificial intelligence (AI) technologies can also play an important role in the fight against child sexual exploitation online by supporting the prevention, detection, and prosecution of such crimes. When employing AI solutions, it is important to ensure compliance with children's rights and their best interests. UNICEF proposes the following principles of child-centred AI:[221]

1. Support children's development and well-being
2. Ensure inclusion of and for children
3. Prioritize fairness and non-discrimination for children
4. Protect children's data and privacy
5. Ensure safety for children
6. Provide transparency, explainability, and accountability for children
7. Empower governments and businesses with knowledge of AI and children's rights
8. Prepare children for present and future developments in AI
9. Create an enabling environment

One of the important steps towards preventing OSEC is being able to **distinguish between children and adults online**. AI can facilitate this process through child recognition systems (including factors such as voice, facial features, and writing style).[222] Voice recognition (or vocal fingerprinting) can also be used in case of audio or video including CSAM to help identify victims and perpetrators.[223] Some facial recognition tools (e.g. ChildSafe) include bin-based classifiers, as well as 3D cameras that identify skeletal features, to distinguish between adults and children. Such technology operates in a three-step manner by, starting by taking human features as input, then ascribing adequate 'bins' to the features and finally making an age estimation to classify the person as a child or as an adult. Such techniques can be useful for large applications to estimate the number of children who use their services to ensure their safety.[224]

Technology is used to **detect nudity** and sexually suggestive photos in combination with the age of the subject, but also sexually suggestive photos. Tariq, Razi, Badillo-Urquiola and Wisniewiski (2019) therefore suggest not to define nudity in 'binary terms', but rather explore what level of nudity could do harm.[225] There are several scales that might be useful to differentiate types of content with the most elaborate one originating from the Combating Paedophile Information Networks in Europe (COPINE) project. The COPINE project developed a ten-point scale ranging from non-sexualized images to sadistic imagery (see Table 3[226]) Nudity detection could, however, mean additional risk for these teens, as their photos might be saved on an additional server that processes these risky images. A solution could be to have a low-powered sensor with added privacy protection measures. Another idea is that technology can intervene when teenagers are about to do something risky online, such as a risky text or a nude photo for instance. If technology alerts the child that this might be dangerous and have consequences, they might rethink the action.[227]

[221] UNICEF (2021) *Policy guidance on AI for children*. Available at: https://www.unicef.org/globalinsight/media/2356/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf (Accessed 25 April 2022).

[222] Nezhad, M. M. and Mehrnezhad, M. (2018) 'A child recognition system based on image selection patterns', *7th Workshop on Socio-Technical Aspects in Security and Trust* (STAST '17), Orlando, United States, 5 December. ACM, New York, United States, pp. 76-81. https://doi.org/10.1145/3167996.3168003.

[223] Bracket Foundation (2019) *Artificial Intelligence – Combating Online Sexual Abuse of Children*. Available at: https://respect.international/wp-content/uploads/2019/11/AI-Combating-online-sexual-abuse-of-children-Bracket-Foundation-2019.pdf (Accessed 25 April 2022).

[224] Basaran, C., Yoon, H. J., Ra, H. K., Son, S. H., Park, T. and Ko, J. G. (2014) 'Classifying children with 3D depth cameras for enabling children's safety applications', *ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '14), Seattle, United States, 13-17 September. ACM, New York, United States, pp. 343-347. https://doi.org/10.1145/2632048.2636074.

[225] Tariq, M. U., Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2019) 'A Review of the Gaps and Opportunities of Nudity and Skin Detection Algorithmic Research for the Purpose of Combatting Adolescent Sexting Behaviors'.

[226] Meridan, H., Thakker, J., Wilson, N. and Boer, D. (2011) 'Assessing the internal structure of the COPINE scale', *Psychology, Crime and Law*, 19(1), pp.21-34. https://doi.org/10.1080/1068316X.2011.598158.

[227] Tariq, M. U., Ghosh, A. K., Badillo-Urquiola, K., Jha, A., Koppal, S. and Wisniewski, P. (2018) 'Designing light filters to detect skin using a low-powered sensor', *SoutheastCon 2018*, St. Petersburg, United States, 19-22 April. IEEE. https://doi.org/10.1109/SECON.2018.8479027.

Table 3. The COPINE Scale[228]

The COPINE Scale	
1. Indicative	Non-erotic and non-sexualized pictures showing children wearing either underwear or swimsuits from either commercial sources or family albums. Pictures of children playing in normal settings, in which the context or organization of pictures by the collector indicates inappropriateness.
2. Nudist	Pictures of naked or semi-naked children in appropriate nudist settings, and from legitimate sources.
3. Erotic	Surreptitiously taken photographs of children in play areas or other safe environments showing either underwear or warying degrees of nakedness.
4. Posing	Deliberately posed pictures of children fully clothed, partially clothed or nakes (where the amount, context and organization suggests sexual interest).
5. Erotic Posing	Deliberately posed pictures of fully, partially clothed or naked children in sexualized or provocative poses.
6. Explicit Erotic Posing	Pictures emphasizing genital areas, where the child is either naked, partially clothed or ffully clothed.
7. Explicit Sexual Activity	Pictures that depict touching, mutual and self-masturbation, oral sex and intercourse by a child, not involving an adult.
8. Assault	Pictures of children being subject to a sexual assault, involving digital touching, involving an adult.
9. Gross Assault	Grossly obscene pictures of sexual assault, involving penetrative sex, masturbation or oral sex, involving an adult.
10.Sadistic/Bestiality	a. Pictures showing a child being tied, bound, beaten, whipped or otherwise subject to something that implies pain. b. Pictures where an animal is involved in some form of sexual behavior with a child.

Source: R v. Oliver (2002) EWCA Crim, case 2766. Available at: *https://vlex.co.uk/vid/r-v-oliver-r-792617673*.

Another example is a proposed algorithm that identifies child users *“based on the patterns users produce when clicking a set of pictures”*.[229] The idea behind such a system is to use cognitive and behavioural aspects to differentiate between adults and children (e.g. children tend to choose pictures that they are more exposed to daily). These are called **behavioural classifiers**. Experiments show that such graphical games can be an effective tool to identify child users,[230] which can help to protect them from being exposed to inappropriate content or contacts that might put them at risk of grooming.

[228] R v. Oliver (2002) EWCA Crim, case 2766. Available at: **https://vlex.co.uk/vid/r-v-oliver-r-792617673** (Accessed 25 April 2022).

[229] Nezhad, M. M. and Mehrnezhad, M. (2018) ‘A child recognition system based on image selection patterns’.

[230] ibid

AI technology can also be used to **identify and stop online grooming** of children at an early stage (e.g. in chat conversations) even if the language is coded or cryptic, as automated tools can help identify potential exploitation patterns.[231] Proposed systems can be based on natural language processing (NLP), for instance a bag of words (BoW) approach, and use sets of classifiers to automatically distinguish conversations linked to child grooming.[232] Chatbots using NLP can be deployed to engage and identify perpetrators online.[233] Safeguarding children online with the support of AI technologies is also promising in the context of detecting promoters of harmful video content and comments on platforms such as YouTube, with a detection rate of 85.7%.[234]

Literature suggests that **while artificial intelligence presents valuable opportunities to increase safety, it will never be 100% successful** due to its statistical nature and the risks of both false negatives and positives. Furthermore, potential biases and consequences need to be acknowledged and addressed in regards to the level of risks (e.g. privacy related) and gains (e.g. contributing to child safety) of applying AI,[235] and stakeholder involvement is crucial. It is also important to ensure that AI systems are prepared for *"live evolving streaming chat conversations along with the evolution of Internet language"*[236] to protect children effectively. In order to concretely support the fight against OSEC through AI, more emphasis needs to be placed on developing tools for prevention and expanding *"the reach of AI solutions to new geographies and new forms of abuse"*.[237] AI-based predictive network analysis seems particularly important in the context of prevention; algorithms use deep learning techniques and multi-sourced data to assess both the likelihood of children falling victim to sexual abuse online and the likelihood of predatory behaviour, and have a deterrent effect by engaging in real time to prevent abuse from happening in the first place.[238] However, over-reliance on algorithms can provide a false sense of safety as AI provides imperfect solutions that still require human moderation and checks.

2.2.6 Behavioural profiling

Technological in-app features should also promote self-regulation based on the age of the child. This means that features for teenagers between 13 to 15 years old could be different to those for older teenagers of 16 or 17 years old.[239] An example of this is making teenagers aware of what strangers can see on their profile[240] or using interface design that provides teenagers with options to protect themselves, or reach out for help when they are at risk.[241] The in-app design could potentially identify behavioural patterns (**behavioural profiling**) that in the past have proven to be risky and alert teens who are following these same patterns.[242] This could make teenagers more aware of what risky behaviour is and what to avoid in the future, for instance.[243] An example of such behavioural profiling is SafetoNet. Their technology scans the messages that children want to send and helps them navigate towards a more appropriate message. The technology can also identify signals

[231] Bracket Foundation (2019) *Artificial Intelligence – Combating Online Sexual Abuse of Children*.

[232] Anderson, P., Zuo, Z., Yang, L. and Qu, Y. (2019) 'An Intelligent Online Grooming Detection System Using AI Technologies', 2019 *IEEE International Conference on Fuzzy Systems* (FUZZ-IEEE), New Orleans, United States, 23-26 June. IEEE, pp.1-6. https://doi.org/10.1109/FUZZ-IEEE.2019.8858973.

[233] Bracket Foundation (2019) *Artificial Intelligence – Combating Online Sexual Abuse of Children*.

[234] Kaushal, R., Saha, S., Bajaj, P. and Kumaraguru, P. (2016) *KidsTube: Detection, characterization and analysis of child unsafe content & promoters on YouTube*. https://doi.org/10.48550/arXiv.1608.05966.

[235] Basaran, C., Yoon, H. J., Ra, H. K., Son, S. H., Park, T. and Ko, J. G. (2014) 'Classifying children with 3D depth cameras for enabling children's safety applications'.

[236] Anderson, P., Zuo, Z., Yang, L. and Qu, Y. (2019) 'An Intelligent Online Grooming Detection System Using AI Technologies'.

[237] Bracket Foundation (2019) *Artificial Intelligence – Combating Online Sexual Abuse of Children*.

[238] ibid

[239] Wisniewski, P., Ghosh, A. K., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Parental Control vs. Teen Self-Regulation: Is there a middle ground for mobile online safety?'.

[240] Wisniewski, P., Jia, H., Wang, N., Zheng, S., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Resilience mitigates the negative effects of adolescent internet addiction and online risk exposure'. *ACM Conference on Human Factors in Computing Systems* (CHI '15), Seoul, South Korea, 18-23 April, pp.4029-4038. https://doi.org/10.1145/2702123.2702240.

[241] Wisniewski, P., Jia, H., Wang, N., Zheng, S., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Resilience mitigates the negative effects of adolescent internet addiction and online risk exposure'; Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors', *ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '15), Vancouver, Canada, 14-18 March. ACM, New York, United States, pp.583-599. https://doi.org/10.1145/2675133.2675287.

[242] Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors'; Nash, V., Adler, J. R., Horvath, M. A. H., Livingstone, S., Marston, C., Owen, G. and Wright, J. (2015) *Identifying the routes by which children view pornography online: implications for future policy-makers seeking to limit viewing*.

[243] Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors'.

such as having low self-confidence or having dark thoughts.[244] Other options for social media apps include using pop-ups to deter possible offenders or to warn teens engaging in risky behaviour.[245]

In a 2020 WeProtect survey, just 37% of participating companies said that they use AI to proactively keep their platform safe through the use of classifiers. An interesting finding is that 50% of respondents use classifiers that are made by other companies, whereas only 26% of companies share the technology and tools that they have developed themselves.[246]

2.2.7 Safety measures used by popular social media apps

Now that the different categories have been explained, this section describes the type of safety measures popular social media platforms implement, with the caveat that only publicly available information was used for this overview and that the lack of transparency on the safety measures used by the industry means that this overview provides an incomplete picture. The summary of the findings is shown in Table 4.

All social media platforms use an age limit of 13, which is in line with child protection acts. The social media platforms verify this age through self-declaration when new users register, for instance by asking for a date of birth. **Meta** and **TikTok** say that they use AI to later verify whether or not this date of birth is actually correct. TikTok is the only app that explicitly states in their community guidelines that an account will be removed if and when they discover that a user is under the age of 13, having claimed to be older.[247] **YouTube** has a special platform for users that are below the age of 13 that had additional safety measures and content control to keep these younger children safe.[248] In the US, TikTok also has a special app for Younger Users especially designed for a younger audience.[249]

Meta and TikTok both differentiate their services for younger and older users. TikTok, for instance, limits the functioning of the app.[250] By default, profiles of younger users aged 13 to 15 are set to private meaning that only people who the child accepted can follow their account and see their content. Additionally, the child's account is not suggested to others, their videos are not downloadable, and direct messaging is not available. Only accepted friends can comment on videos.[251] Instagram provides new users under the age of 18 with information about the differences between private and public, with private selected by default. Meta says that they understand that younger people also see the benefits of having public accounts and therefore allow them to choose. On **Instagram**, adults cannot contact users under the age of 18 if they do not follow them. Meta also has the option of pop-up messages that warn teens when they are engaging in risky conversations.[252]

All social media platforms have the option for in-app reporting, after which the report is viewed and a decision is made about the content. The content could, for instance, be removed. Meta, TikTok and YouTube all have technology and classifiers in place that detect content violations and remove the content, thus taking a proactive approach. Meta and YouTube both have a collaboration with the National Center for Missing and Exploited Children (NCMEC), where they report content related to the exploitation of children. Instagram and Facebook seem to be the only platforms that use messaging to discourage risky behaviour.

[244] Donaldson, S., Davidson, J. and Aiken, M. (2021) Safer technology, safer users: *The UK as a world-leader in Safety Tech*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/974414/Safer_technology__safer_users-_The_UK_as_a_world-leader_in_Safety_Tech_V2.pdf (Accessed 27 April 2022).

[245] WeProtect Global Alliance (2021) *Global Threat Assessment 2021 – Working together to end the sexual abuse of children online*.

[246] ibid

[247] TikTok (n.d.) *New User Guide*. Available at: https://www.tiktok.com/safety/en/new-user-guide/ (Accessed 25 April 2022).

[248] The Verge (2021). The child safety problem on platforms is worse than we knew Available at: https://www.theverge.com/2021/5/12/22432863/child-safety-platforms-thorn-report-snap-facebook-youtube-tiktok (Accessed 25 April 2022).

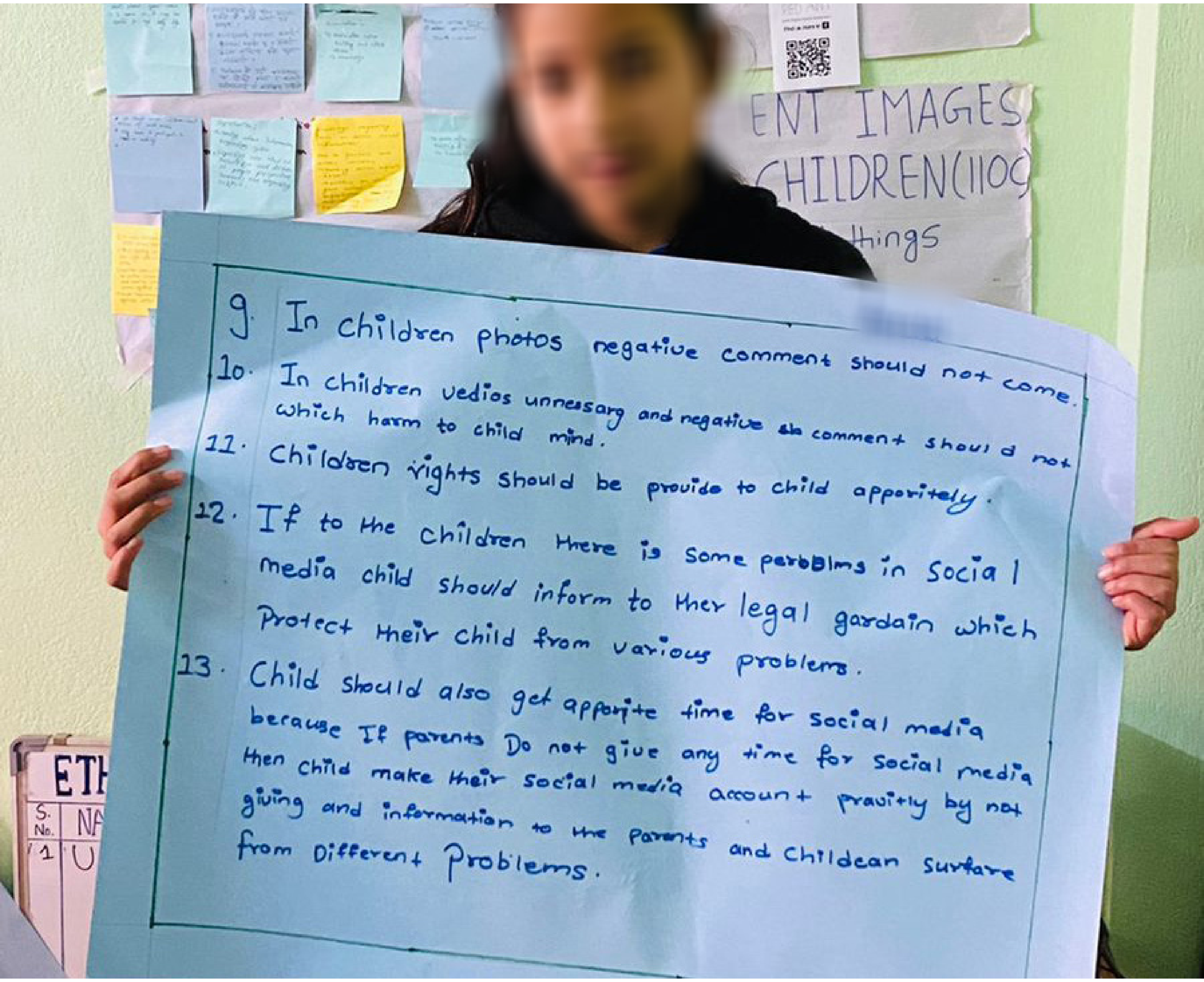
[249] TikTok (n.d.) *Community Guidelines – Minor safety*. Available at: https://www.tiktok.com/community-guidelines?lang=en#31 (Accessed 25 April 2022).

[250] TikTok (2021) *Legal – Privacy Policy*. Available at: https://www.tiktok.com/legal/privacy-policy-eea?lang=en (Accessed 25 April 2022).

[251] TikTok (n.d.) *Teen privacy and safety settings*. Available at: https://support.tiktok.com/en/account-and-privacy/account-privacy-settings/privacy-and-safety-settings-for-users-under-age-18 (Accessed 25 April 2022).

[252] Meta (2021) *Instagram: Continuing to Make Instagram Safer for the Youngest Members of Our Community*. 16 March. Available at: https://about.fb.com/news/2021/03/continuing-to-make-instagram-safer-for-the-youngest-members-of-our-community/ (Accessed 25 April 2022).

In comparison to the other apps, **Snapchat** appears to be using fewer measures to keep children safe. No source was identified indicating whether Snapchat makes use of classifiers or deterrence messages. Snapchat removes content that is reported and that is proven to be against their community guidelines. If the platform finds that a certain account has violated the rules, they *"may remove the offending content, terminate or limit the visibility of your account, and/or notify law enforcement"*.[253] These are all possibilities, but there does not seem to be a strict policy concerning these topics. Snapchat thus appears to be taking a more reactive approach. Nevertheless, Snapchat is designed with a stronger privacy focus. By default, only friends added by the user can view a user's content or contact the user. In addition, content is deleted by default and if someone takes a screenshot of a piece of content, the user who posted it is notified.[254]



9. In children photos negative comment should not come.

10. In children vedios unnesarg and negative ~~sh~~ comment shoul d not which harm to child mind.

11. Children rights should be provide to child apporitely.

12. If to the children there is some perbblms in socia l media child should inform to ther legal gardain which Protect their child from various problems.

13. Child should also get apporite time for social media because If parents Do not give any time for social media then child make their social media account pravitly by not giving and information to the parents and childean surfare from Different Problems.

Source: Focus group in Nepal

[253] Snap Inc. (2022) *Community Guidelines*. Available at: https://snap.com/en-US/community-guidelines (Accessed 25 April 2022).

[254] Snap Inc. (n.d.) *Privacy settings*. Available at: https://support.snapchat.com/en-GB/a/privacy-settings2 (Accessed 7 May 2022).

Table 4. Overview of safety measures that popular social media apps appear to be using

	Minimum age	Age verification	Automatic settings for children	Use of classifiers	Reporting and deleting of material	Deterrence and warnings
Facebook and Instagram[255]	13	Deterrence and warnings	Yes, shielding child users from messages from unknown adults and preventing adults from searching for children	Yes	In-app reporting Deletion of material Reporting to the NCMEC	Yes
Snapchat[256]	13	Self-declaration	No, but general privacy by default for all users	Unknown	In-app reporting Deletion of material	Unknown
TikTok[257]	13, with TikTok for Younger Users available in the US for children under 13	Self-declaration with possibility of termination of profile if found to have lied about being 13 or older	Yes, the most safety settings for children, including pin-protected parental controls	Yes	In-app reporting Deletion of material	Unknown
YouTube[258]	13, with YouTube Kids available for children under 13 since 2015	Self-declaration	In YouTube Kids, yes. Plus the option to have a supervised account	Yes	In-app reporting Removal of material Reporting to NCMEC	Unknown

[255] Meta (2021) *Instagram: Asking People for Their Birthday on Instagram*. 30 August. Available at: https://about.fb.com/news/2021/08/asking-people-for-their-birthday-on-instagram/ (Accessed 27 April 2022); Meta (2018) Facebook: *New Technology to Fight Child Exploitation*. 24 October. Available at: https://about.fb.com/news/2018/10/fighting-child-exploitation/ (Accessed 27 April 2022); Meta Transparency Center (2022) Child Sexual Exploitation, Abuse and Nudity; WeProtect Global Alliance (2021) *Global Threat Assessment 2021 – Working together to end the sexual abuse of children online*; The Verge (2021) The child safety problem on platforms is worse than we knew, 12 May. Available at: https://www.theverge.com/2021/5/12/22432863 (Accessed 8 May 2022).

[256] Snap Inc. (n.d.) *Snapchat Safety Center*. Available at: https://snap.com/en-US/safety/safety-center (Accessed 27 April 2022); Snap Inc. (n.d.) *Privacy settings*.

[257] Nealon, L. (2021) 'TikTok Takes Fighting Sexploitation Seriously', *National Center on Sexual Exploitation*. 14 July. Available at: https://endsexualexploitation.org/articles/tiktok-takes-fighting-sexploitation-seriously/ (Accessed 27 April 2022); TikTok (n.d.) *Community Guidelines – Minor safety*; WeProtect Global Alliance (2021) *Global Threat Assessment 2021 – Working together to end the sexual abuse of children online*; TikTok (n.d.) New User Guide.

[258] The Verge (2021) *The child safety problem on platforms is worse than we knew*; YouTube (n.d.) *Community Guidelines*. Available at: https://www.youtube.com/intl/en-GB/howyoutubeworks/policies/community-guidelines/#enforcing-community-guidelines (Accessed 25 April 2022); YouTube (n.d.) *Child safety policy*. Available at: https://support.google.com/youtube/answer/2801999?hl=en&ref_topic=9282679 (Accessed 25 April 2022).

2.3 CHALLENGES TO IMPLEMENTING CHILD SAFETY DESIGNS

Developing and implementing effective safety by design raises a number of challenges, some of which are discussed in further detail below. The particularity of the challenges is that they take the form of dilemmas. Ensuring the safety of children by protecting them from harms risks encroaching on their right to actively take part in the online environment. Overly protective approaches also prevent children from learning to deal with risk. Restricting sexual content and interactions make platforms safer, but this limits teenagers' capacity to explore their own sexuality. Social media platforms promote a culture of sharing, which has many benefits, yet leads to privacy issues and risks. There is no single solution that solves these dilemmas, yet any potential safety by design solution must account for them and must find ways to ensure a balance between protection and freedom; between sharing and privacy; and between safety and profit.

2.3.1 Freedom of expression

The child's right to freedom of expression is enshrined in Article 13 of the UNCRC, which includes *"the freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of the child's choice"*.[259] The digital environment provides endless possibilities for children to express themselves and to search and retrieve information from the internet, especially heightened by extreme circumstances such as the COVID-19 pandemic. As a legal principle, **any restrictions to the child's freedom of expression in the digital environment**, such as filters and other safety measures, should be established in law, should be deemed necessary and proportionate, and should be transparent and well communicated to children in age-appropriate language.[260] The UN CRC, in is General comment No. 25 has called for States to adopt guidelines, standards, and codes that enable children to safely access diverse content. This should be done while recognising children's rights to information and freedom of expression, while protecting them from harmful material in accordance with their rights and evolving capacities.[261]

Automated filtering processes should not interfere with children's ability to form and express their opinions in the digital environment. Automated filtering enables platforms to sort through the vast amount of material uploaded on their platforms by classifying, demoting, or excluding (user-generated) materials that are (potentially) illegal and abusive. While useful, automated filtering presents the risk of excluding legal content, infringing on the freedom of expression. Digital service providers must ensure necessary and proportionate content moderation rules and safety-oriented technologies, to filter out the flow of harmful and illegal material, while leaving intact all other material protected under the right to freedom of expression.

While protecting children from OSEC should be a top priority, it should not be a reason to overtly police, censor, or surveil children and young people online through digitised means and algorithms.[262] **There is a fine line between protecting children's freedom of expression and protecting them from harm**. Teenagers are often unaware, or rather underestimate the risks of their online behaviour. For instance, 'sexting', the act of sending sexually suggestive or explicit messages or photographs through peer-to-peer communication platforms, is a phenomenon becoming increasingly popular among young people.[263] In these situations, promoting free expression and non-judgement is important in order to protect children and youth from the harms and risks that come along with sexting. Protecting children from OSEC does not necessarily mean overtly limiting their freedom, but rather it should be a **learning, supporting, open, and continuous process** between children, parents, legislators, and online service providers. Children prefer to have features that allow them to manage online risk and empower them to protect themselves, instead of parental control approaches which may limit the child's freedom of expression.[264] Being overly cautious and taking a risk-averse stance to online safety may

[259] United Nations (1990) ***United Nations Convention on the Rights of the Child***, Article 13. Available at: **https://www.ohchr.org/sites/default/files/crc.pdf** (Accessed 27 April 2022).

[260] Smirnova, S., Livingstone S. and Stoilova, M. (2021) ***Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls***.

[261] UN CRC (2021) ***General comment No. 25 (2021) on children's rights in relation to the digital environment***.

[262] Razi, A., Kim, S., Alsoubai, A., Stringhini, H., Solorio, T., De Choudhury, M. and Wisniewski, P. (2021) 'A Human-Centered Systematic Literature Review of the Computational Approaches for Online Sexual Risk Detection'.

[263] INHOPE (2021) ***What is sexting?*** Available at: **https://inhope.org/EN/articles/what-is-sexting?** (Accessed 27 April 2022).

[264] Wisniewski, P., Ghosh, A. K., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Parental Control vs. Teen Self-Regulation: Is there a middle ground for mobile online safety?'.

hamper children's developmental growth.[265] Children need to have the space to grow and develop resilience, especially within the digital world. 'Zero risk' design should not be the goal.

2.3.2 Privacy

A child's right to privacy is embedded in various children's rights conventions, such as in Article 16 of the UNCRC, which states that a child may not be subjected to any arbitrary or unlawful interference with their privacy, family, home, or correspondence. The child, under Article 16, may not be subjected to unlawful attacks on their honour and reputation. States have the primary obligation to protect an individual's privacy from unlawful and arbitrary interference. The UN Special Rapporteur on the Right to Privacy has introduced a **four-fold test** to determine if an infringement of the right to privacy is legitimate. According to this test, a legitimate interference should be:

1. non-arbitrary and provided for by law;
2. necessary in a democratic society;
3. for the purpose of national security or public safety, public order, the protection of public health or morals or the protection of the rights and freedoms of others; and
4. proportionate to the threat or risk being managed.[266]

Within the European framework, the right to privacy is embedded in Article 8 of the European Convention on Human Rights (ECHR) and Article 7 of the EU Charter of Fundamental Rights, and is reinforced and implemented through legal acts such as the GDPR. More specifically, the GDPR states that children, defined as anyone below the age of 16, require specific protection with regard to their personal data. Children may be less aware of the risks, consequences and safeguards to their rights in relation to the processing of their personal data.

There is currently a tension between the digital privacy of the child, the child's right to freedom of expression, and the inherent need for child protection against online sexual exploitation and abuse.[267] A gap that continues to exist in research is children's own perception of their right to privacy, paying special attention to what privacy in the digital world means for children.[268] Understanding what privacy is from the child's perspective is an important first step in balancing their privacy needs with the need for protection.

Detection of CSAM or potential OSEC through the use of artificial intelligence may have implications on the child's right to privacy, among other rights, given the child's evolving capacity and heightened vulnerabilities.[269] As mentioned above, the child's right to privacy includes the protection from any unlawful interference with the child's correspondence, and yet automated detection of content or conversations for grooming already constitutes an interference with the child's correspondence. Privacy advocates have raised strong concerns about the automated detection of CSAM, yet children's right to privacy continues to be violated when CSAM is not taken down or when material depicting them is continuously shared without the child's consent. Thus, it is important for legislators, when adopting the regulation of AI and other technologies used to detect CSAM or grooming, to carefully assess the necessity and proportionality of using such technology. This need is heightened by the fact that children are more vulnerable to intrusions of their privacy.[270] Transparency between technology providers, social media platforms, legislators, and governments is key to ensuring that these tools are used for

[265] Pinter, A. T., Wisniewski, P., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Adolescent Online Safety: Moving Beyond Formative Evaluations to Designing Solutions for the Future'. *Conference on Interaction Design and Children* (IDC '17), Stanford, United States, 27-30 June. ACM, New York, United States, pp.352-357. https://doi.org/10.1145/3078072.3079722.

[266] UNHCR (2018) *Report of the Special Rapporteur on the Right to Privacy*, paragraph 55. Available at: https://digitallibrary.un.org/record/1656178/files/A_HRC_37_62-EN.pdf (Accessed 27 April 2022).

[267] Stoilova, M., Livingstone, S. and Khazbak, R. (2021) *Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children's internet use and outcomes*. Available at: https://www.unicef-irc.org/publications/pdf/Investigating-Risks-and-Opportunities-for-Children-in-a-Digital-World.pdf (Accessed 27 April 2022).

[268] ibid

[269] UNICEF Innovation, Human Rights Center, UC Berkley (2019) *Executive Summary: Artificial Intelligence and Children's Rights*. Available at: https://www.unicef.org/innovation/media/10726/file/Executive%20Summary:%20Memorandum%20on%20Artificial%20Intelligence%20and%20Child%20Rights.pdf (Accessed 27 April 2022).

[270] UNICEF Innovation, Human Rights Center, UC Berkley (2019) *Memorandum on Artificial Intelligence and Child Rights*. Available at: https://www.unicef.org/innovation/media/10501/file/Memorandum%20on%20Artificial%20Intelligence%20and%20Child%20Rights.pdf (Accessed 27 April 2022).

no more than their sole purposes of protecting children's privacy and right to be free from violence. In addition to safety by design, **privacy by design** is an important principle that tech companies and legislators must keep in mind. It has been identified that many protective features such as age assurance measures do not respect children's right to privacy, and privacy by default or by design is rarely implemented.[271]

Another important angle to look at is the tension that arises between children and their parents when it comes to online privacy. As children continue to seek independence in the digital world, there is often a blur when it comes to the role of parents, legal guardians, or teachers in interfering with the child's online presence. Given children's evolving capacity, it is important to give teenagers in particular the space to develop resilience against privacy risks, while still avoiding as many risks as possible. An important principle to remember when designing such measures is that **children should be treated as independent rights holders**. Restricting children's internet access may not always limit online risk. In contrast, it may result in limiting the opportunities that are open to the child in the digital world, such as learning, creating relationships, and expressing themselves.

Children value their independence in terms of internet use.[272] **Measures to protect them must not promote an age-inappropriate or undesirable degree of parental surveillance.** Therefore, it is important to promote an environment of communication and respect when it comes to parents, guardians and their mediation of the child's social media use. It has been identified that conversations between children and their parents and teachers about privacy positively affect children's privacy behaviours, such as having private profiles online. However, these conversations decrease as the child gets older.[273] Resources and information on privacy matters and being a good digital citizen must be made available to parents and families, such as on how to have conversations with their children about CSAM and OSEC, as an alternative to a strict monitoring approach.

2.3.3 Companies' buy-in and profits

When designing their products, companies set out multiple, often competing, objectives. They want their platforms to be attractive to (new) users and facilitate exchange of images, text, and videos. They often aim to design safe platforms, yet promote content that will create reactions, using nudges that encourage interactions and keep users active on their platforms, all of which can lead to risks for children's online safety. **Dark patterns** nudge users into certain actions without them noticing. These patterns are widely used on platforms and may result in making the online space less safe for children. For example, Facebook has been found to design their platforms in a way that nudges users away from stricter privacy choices and makes it complex and time-consuming to opt for stronger privacy settings.[274] Other nudging techniques include encouraging young people to add strangers as 'friends' or 'followers' and location tracking used by certain friend suggestion systems.[275] Recent whistle-blower revelations show that when facing competing priorities, **platforms often choose to prioritise profit over online safety**.[276]

Research on child privacy in online apps shows that while developers acknowledge the importance of children's best interests and believe that apps for children should be designed differently than apps for adults (in terms of age-appropriateness and developmental needs) they tend to compromise on these beliefs for several reasons:

[271] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) U*nderstanding of user needs and problems: a rapid evidence review of age assurance and parental controls*, p.6.

[272] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls*, p.4.

[273] Stoilova, M., Livingstone, S. and Khazbak, R. (2021) *Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children's internet use and outcomes*.

[274] Norwegian Consumer Council (2018) *Deceived by design: How tech companies use dark patterns to discourage us from exercising our rights to privacy*.

[275] 5Rights Foundation (n.d.) *Risky-by-Design*. Available at: https://www.riskyby.design/friend-suggestions (Accessed 18 April 2022).

[276] Hao, K. (2021) 'The Facebook whistleblower says its algorithms are dangerous. Here's why.', *MIT Technology Review*, 5 October. Available at: https://www.technologyreview.com/2021/10/05/1036519/facebook-whistleblower-frances-haugen-algorithms/ (Accessed 27 April 2022).

1. **Third-party libraries** are used to collect sensitive data about online users, including children, which is then used for data profiling (e.g. for targeted advertising). These libraries contain sensitive information about children, such as their location and can *"track call logs, browser history, and contact information for the purpose of targeted advertisements"*.[277] This lack of a child-focused approach is worrying, as it has been shown that children as young as 8 years of age tend to create accounts on social media without understanding the implications of agreeing for their data to be sold to third parties.[278]
2. **Making apps financially profitable** without relying on advertising is difficult, as more privacy-friendly business models (e.g. premium apps) are rendered unsustainable due to market pressure. Companies depend on revenue from targeted advertising linked to the use of third-party libraries (e.g. Google Analytics, Firebase and Facebook APIs) that rely on large-scale user data profiling.[279]
3. **A lack of awareness** and clear guidelines on designing for children to ensure safety and meet their developmental needs means that developers rely on the examples set by market leaders (e.g. Google, whose processes lack transparency) without necessarily considering children's best interests.[280]

Barriers hindering the adoption of child privacy-friendly apps are of both an economic and social-technical nature, including the targeted advertising revenue model.

The privacy changes seen in recent years (e.g. GDPR) are placing more responsibility on developers to create appropriate apps for children. At the same time, data protection frameworks fail to address these barriers and adequately support app developers.[281]

Other research points to how a **lack of transparency** about the consequences of choosing particular third-party libraries impacts app design, as choice framing of data practices influences developers' decisions. As has been discussed, big companies tend to set an example (e.g. in terms of advertising networks) for smaller or independent developing enterprises, which once again highlights the need to ensure that the big players comply with safety practices.[282]

Results of a survey conducted among software developers in North America show that the issues stem from the **lack of organisational/process support** to handle security throughout development tasks.[283] Interestingly, the motivators for developers to consider security in app development have been shown to be strongly self-driven (e.g. feeling responsible for user protection, understanding the implications of security, caring about the reputation of the company, or identifying with the importance of security in principle) while financial rewards were reported as less motivating. It would therefore be important to support these intrinsic motivations while ensuring a formal process is in place to support developers and enhance awareness of security code-analysis tools.[284] It is also important to note that developers have also reported feeling unable to, and not responsible for addressing consumer risks.[285]

Software developers have a crucial role to play in ensuring app safety. At the same time, they seem to *"find themselves making trade-offs in protecting children and sustaining their business"*.[286]

[277] Ekambaranathan, A., Zhao, J. and Van Kleek, M. (2021) '"Money makes the world go around": Identifying Barriers to Better Privacy in Children's Apps From Developers' Perspectives'. ***CHI Conference on Human Factors in Computing Systems*** (CHI '21), Yokohama, Japan. 8-13 May. ACM, New York, United States, Article 46, pp.1-15. **https://doi.org/10.1145/3411764.3445599**.

[278] Nezhad, M. M. and Mehrnezhad, M. (2018) 'A child recognition system based on image selection patterns'.

[279] Ekambaranathan, A., Zhao, J. and Van Kleek, M. (2021) '"Money makes the world go around": Identifying Barriers to Better Privacy in Children's Apps From Developers' Perspectives'.

[280] ibid

[281] ibid

[282] Tahaei, M., Frik, A. and Vaniea, K. (2021) 'Deciding on Personalized Ads: Nudging Developers About User Privacy'. ***USENIX Symposium on Usable Privacy and Security*** (SOUPS 2021), Virtual Event, 8-10 August. **https://doi.org/10.7488/ds/3045**.

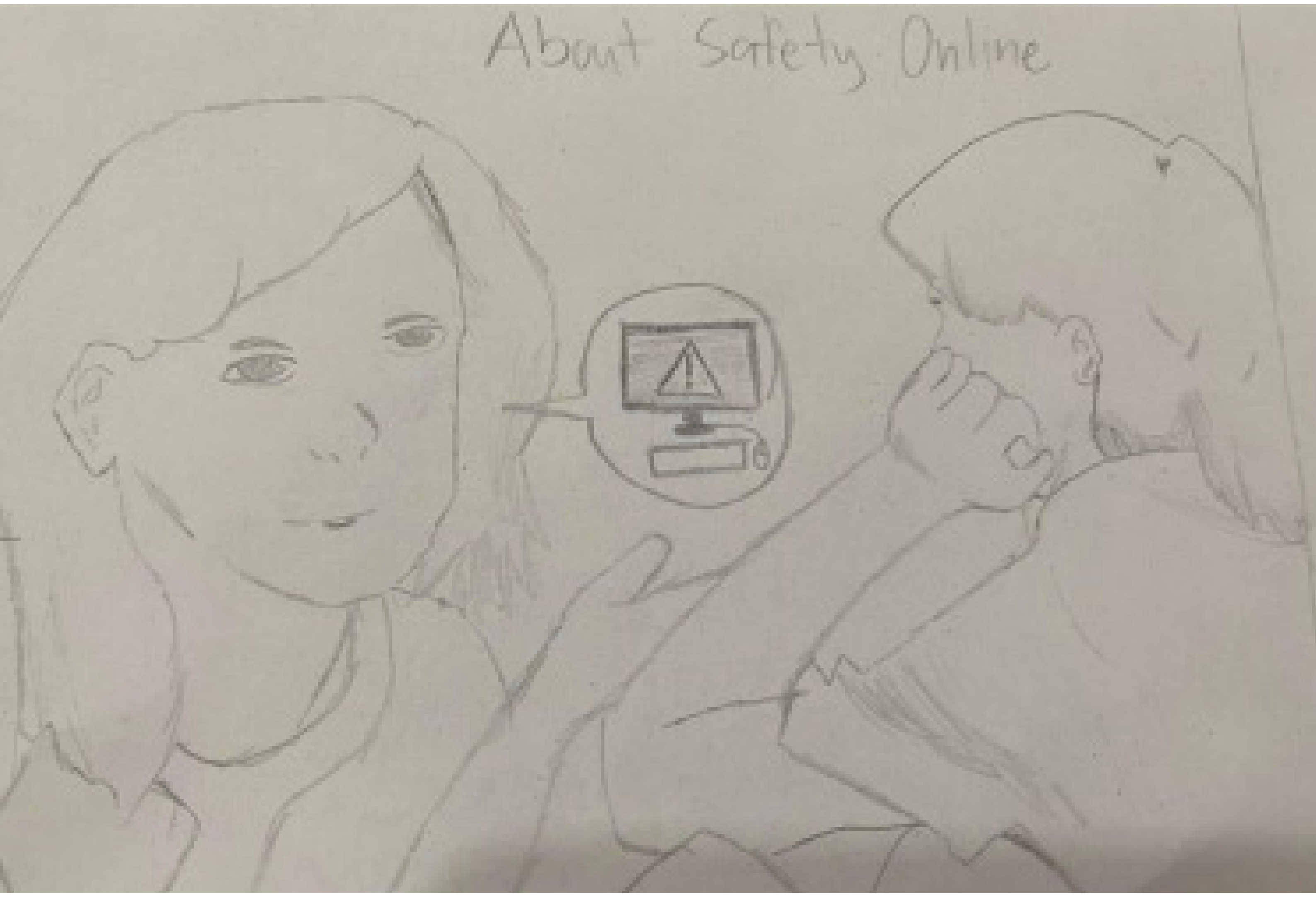
[283] Assal, H. and Chiasson, C. (2019) '"Think secure from the beginning": A Survey with Software Developers'. ***CHI Conference on Human Factors in Computing Systems Proceedings*** (CHI '19). Glasgow, United Kingdom, 4-9 May. ACM, New York, United States. 13 pages. **https://doi.org/10.1145/3290605.3300519**.

[284] ibid

[285] Mhaidli, A.H., Zou, Y. and Schaub, F. (2019) '"We Can't Live Without Them!" App Developers' Adoption of Ad Networks and Their Considerations of Consumer Risks' ***Fifteenth USENIX Symposium on Usable Privacy and Security***. Santa Clara, United States, 12-13 August, pp.225-244. Available at: **https://www.usenix.org/system/files/soups2019-mhaidli.pdf** (Accessed 25 April 2022).

[286] Ekambaranathan, A., Zhao, J. and Van Kleek, M. (2022) *How Can We Design Privacy-Friendly Apps for Children? Using a Research Through Design Process to Understand Developers' Needs and Challenges*.

Research conducted among Android app developers on their approach to security suggests that they do tend to make security updates when potential damage to stakeholders is understood or when a security expert is involved.[287] However, only 17% of respondents reported being supported by security experts. Assurance techniques are used by less than 30% of app developers, showing the need to strengthen support for these developers to incorporate security measures despite the lack of security professionals. In addition, the findings suggest a limited impact of GDPR compliance rules.[288] One of the proposed solutions would be to involve companies in supporting their developers by strengthening collaboration between developers and experts to enable them to improve their coding by putting more emphasis on safety.[289]



Source: Focus group in the Philippines

[287] Weir, C., Hermann, B. and Fahl, S. (2020) *From Needs to Actions to Secure Apps? The Effect of Requirements and Developer Practices on App Security*. Available at: https://eprints.lancs.ac.uk/id/eprint/142148 (Accessed 25 April 2022).

[288] Weir, C., Hermann, B. and Fahl, S. (2020) *From Needs to Actions to Secure Apps? The Effect of Requirements and Developer Practices on App Security*.

[289] Assal, H. and Chiasson, C. (2019) '"Think secure from the beginning": A Survey with Software Developers'.

3. Child safety designs that work against OSEC

KEY MESSAGES

Key reflections for making children safe

- **Children are both part of the problem and part of the solution**: they are risk-takers and overestimate their ability to cope with risks; yet they are part of the solution as it is more effective to involve them in assessing risks online and guiding them in dealing with the risks they face
- **Overly restrictive parental control** and fear-based abstinence approaches may reassure parents but are not effective as children will prioritise social benefits over risks
- Helping teenagers deal with online risks is an important part of their development

Technical solutions

- Parental monitoring and restriction should be limited to children aged 3 to 12 years old, while **privacy-preserving parental control** apps providing high-level rather than detailed information are a better fit for teenagers when they recognise their need for autonomy and agency
- For adolescents (13 to 17 years old), safety features should empower them to protect themselves when using technologies through **self-regulation** and self-monitoring, and promote **shared responsibility**, trust and communication between parents and children
- **Peer support platforms** can provide a safe space for children to deal with risks if they are designed to ensure anonymity and safety while moderated by professionals
- **Intelligent privacy by default** should become the industry standard to ensure children's accounts are set to 'friends only' by default, and geo-tagging cannot be done without permission
- Platforms should provide young people with **retroactive privacy features**, such as the ability to untag, delete, block, and report inappropriate content
- Age verification and assurance can be strengthened by using officially provided, automatically generated, user reported, and third party data but privacy must be kept in mind
- Evidence-based computational **risk detection** combined with **risk mitigation** strategies can help identify risks and prompt children to respond to those risks

Solutions designed by children

- Children proposed **feasible ways** of strengthening the design of platforms, adding tweaks and especially asking for **more visibility** of rules, regulations and reporting measures
- Use **popular media and well-known people** such as influencers to spread awareness about child safety online in a more attractive and engaging way
- Make the design of platforms **different for certain age** groups, with safety features adapted to that age
- **Platforms should play a bigger role** in keeping children safe through monitoring, intervening via pop-up messages, and making reporting easier by providing children with many options in a clearly visible way and by punishing violators

3.1 KEEPING CHILDREN SAFE ONLINE

Before looking at the solutions identified by the literature and by children themselves, this Section provides an understanding of the key elements that keep children safe online. This aims to bring context to the solutions presented further in this chapter.

3.1.1 Over-shielding children does not work

To protect children from all these risks, various studies recommend an abstinence approach, which advocates for measures that will lead to less risk exposure, such as discouragement of disclosing personal information[290] or parental control software that filters materials online.[291] Surveillance and tracking perpetuates paranoia and fear for both parents and children. For instance, research has shown that mobile-based location tracking had the potential to undermine trust.[292] Studies analysing the online reviews of parental control apps report that restrictive apps tended to nudged parents to set up **overly restrictive controls and surveillance, leaving children with a feeling of oppression and invasion of privacy**, affecting their social lives and activities, as well as negatively affecting their relationship with their parents.[293] Children reported that some features prevented them from carrying out everyday tasks such as homework and went beyond restricting them from harmful activities.

"Websites should be usable, not only safe" (Focus group discussions, 12-year-old girl, Estonia)

Children generally agreed with apps restricting access to inappropriate behaviour (e.g. accessing pornographic websites) and features that made them feel safer, but disagreed with privacy-invasive approaches. Similarly, children liked apps that helped establish a trusting relationship with their parents in relation to online use, as well as features rewarding positive behaviours.[294]

While abstinence-based or restrictive approaches can be effective for younger children (2 to 12 years old),[295] they are not realistic for teenagers who want to connect with peers online. Research suggests that teenagers prioritise the social benefits of online engagement over the risks.[296] Children could therefore circumvent these parental control measures as they want to spend time online. This undermines the effectiveness of parental control measures.[297]

Research shows **that supportive parent-children relationships and positive experiences at schools are protective factors against online risks, while parental restrictions on children's online activity do not lead to reduced risks.**[298]

[290] Pinter, A. T., Wisniewski, P., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Adolescent Online Safety: Moving Beyond Formative Evaluations to Designing Solutions for the Future'. *Conference on Interaction Design and Children* (IDC '17), Stanford, United States, 27-30 June. ACM, New York, United States, pp.352-357. **https://doi.org/10.1145/3078072.3079722**.

[291] Wisniewski, P., Ghosh, A. K., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Parental Control vs. Teen Self-Regulation: Is there a middle ground for mobile online safety?', *2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. Portland, United States. 25 February – 1 March. ACM, New York, United States. **https://doi.org/10.1145/2998181.2998352**.

[292] Ghosh, A. K., Badillo-Urquiola, K., Guha, S., LaViola Jr., J. J. and Wisniewski, P. (2018) 'Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control'. *CHI Conference on Human Factors in Computing Systems* (CHI '18), Montréal, Canada, 21-26 April, pp.1-14. **https://doi.org/10.1145/3173574.3173698**.

[293] Ghosh, A. K., Badillo-Urquiola, K., Guha, S., LaViola Jr., J. J. and Wisniewski, P. (2018) 'Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control'; Wang, G., Zhao, J., Van Kleek, M. and Shadbol, N. (2021) 'Protection or punishment? relating the design space of parental control apps and perceptions about them to support parenting for online safety'. *Proceedings of the ACM on Human-Computer Interaction,* 5(CSCW2), 343, pp.1-26 **https://doi.org/10.1145/3476084**.

[294] Ghosh, A. K., Badillo-Urquiola, K., Guha, S., LaViola Jr., J. J. and Wisniewski, P. (2018) 'Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control'.

[295] Wisniewski, P., Ghosh, A. K., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Parental Control vs. Teen Self-Regulation: Is there a middle ground for mobile online safety?'.

[296] Pinter, A. T., Wisniewski, P., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Adolescent Online Safety: Moving Beyond Formative Evaluations to Designing Solutions for the Future'.

[297] Cabello-Hutt, T., Cabello, P. and Claro, M. (2018) 'Online opportunities and risks for children and adolescents: The role of digital skills, age, gender and parental mediation in Brazil', *New Media & Society*. 20(7), pp.2411-2431. **https://doi.org/10.1177/1461444817724168**.

[298] Davis, K. and Koepke, L. (2015) 'Risk and protective factors associated with cyberbullying: Are relationships or rules more protective?'. *Learning, Media and Technology*, 41(4), pp.521-545. **https://doi.org/10.1080/17439884.2014.994219**.

The internet is also an ambivalent place, which make it hard for teenagers not to encounter risks; to be able to socially connect online, teenagers need to disclose personal information and engage with others, and it is therefore not unthinkable that they come across pornographic images or send compromising images.[299]

Children, especially teenagers, are constructing their social identity online. They are learning how to disclose information and interact with others online and how this impacts their social connections and forms impressions. Parental control and monitoring of children's online activities means that children are learning how to navigate their online social lives under the watchful eye of their parents, which can create tensions between parents and children. Apps allowing parents to exert too much control over children's online activities tend not to support active parental mediation and undermine children's relationships with their parents.

3.1.2 Children need to learn how to deal with online risks

Shielding teenagers too much could also **hamper their development** as they will not learn from, and how to deal with (online) risks.[300] Restricted internet use could also result in issues such as decreased psychological well-being, low self-esteem, and feeling left out,[301] as it does not fix the underlying needs that teenagers have for social interaction.[302] Existing technology, such as parental control apps, are thus more aimed at comforting parents in their worries about the online risks their children face instead of focusing on the needs of teenagers.[303] This raises the concern that teenagers might find different routes to satisfying what they are looking for on the internet. They could resort to using more hidden ways, which are harder to monitor.[304] Monitoring of teenager's social media accounts could also be circumvented, for instance, by having two social media accounts: one that is suitable for parents viewing and a second account where they can behave as they actually want to.[305]

While **parental controls** focused on a surveillance model may ease the anxieties of parents, they are insufficient and potentially harmful for child-parent trust, communication, and overall relationships.[306] A survey in which children expressed their negative opinions about restrictive and invasive parental control applications, pleading for more respectful and moderate solutions, confirmed this.[307] This is why research suggests putting emphasis on flexibility, **fostering mediation**, and **positive involvement** in children's online activity instead of implementing strict, technical surveillance. After all, excessive restrictions may unnecessarily limit children's access to digital rights and their ability to **develop resilience** against online risks. Furthermore, such restrictions may be counterproductive by *"making prohibited behaviours or content more appealing"*.[308] Controlling approaches also lead children to shy away from reporting online harms and risks to parents for fear

[299] Livingstone, S. and Smith, P.K. (2014) 'Annual Research Review: Harms experienced by child users of online and mobile technologies: the nature, prevalence and management of sexual and aggressive risks in the digital age', *The Journal of Child Psychology and Psychiatry*, 55(6), pp.635-54. https://doi.org/10.1111/jcpp.12197.

[300] Pinter, A. T., Wisniewski, P., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Adolescent Online Safety: Moving Beyond Formative Evaluations to Designing Solutions for the Future'.

[301] Erickson, L. B., Wisniewski, P., Xu, H., Carroll, J. M., Rosson, M. B. and Perkins, D. F. (2016) 'The boundaries between: Parental involvement in a teen's online world', *Journal of the Association for Information Science and Technology,* 67(6), pp.1384-1403. https://doi.org/10.1002/asi.23450.

[302] Wisniewski, P., Jia, H., Wang, N., Zheng, S., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Resilience mitigates the negative effects of adolescent internet addiction and online risk exposure'. *ACM Conference on Human Factors in Computing Systems* (CHI '15), Seoul, South Korea, 18-23 April, pp.4029-4038. https://doi.org/10.1145/2702123.2702240.

[303] Finkelhor, D., Jones, L. and Mitchell, K. (2021) 'Teaching Privacy: A flawed strategy for children's online safety'. *Child Abuse & Neglect*, 117. https://doi.org/10.1016/j.chiabu.2021.105064; Wisniewski, P., Ghosh, A. K., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Parental Control vs. Teen Self-Regulation: Is there a middle ground for mobile online safety?'.

[304] Nash, V., Adler, J. R., Horvath, M. A. H., Livingstone, S., Marston, C., Owen, G. and Wright, J. (2015) *Identifying the routes by which children view pornography online: implications for future policy-makers seeking to limit viewing*. Available at: http://eprints.lse.ac.uk/65450/1/__lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_Livingstone%2C%20S_Identifying%20the%20routes_Livingstone_Identifying%20the%20routes_2016.pdf (Accessed 22 April 2022).

[305] Erickson, L. B., Wisniewski, P., Xu, H., Carroll, J. M., Rosson, M. B. and Perkins, D. F. (2016) 'The boundaries between: Parental involvement in a teen's online world'.

[306] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls*. Available at: http://eprints.lse.ac.uk/112559/ (Accessed 22 April 2022).

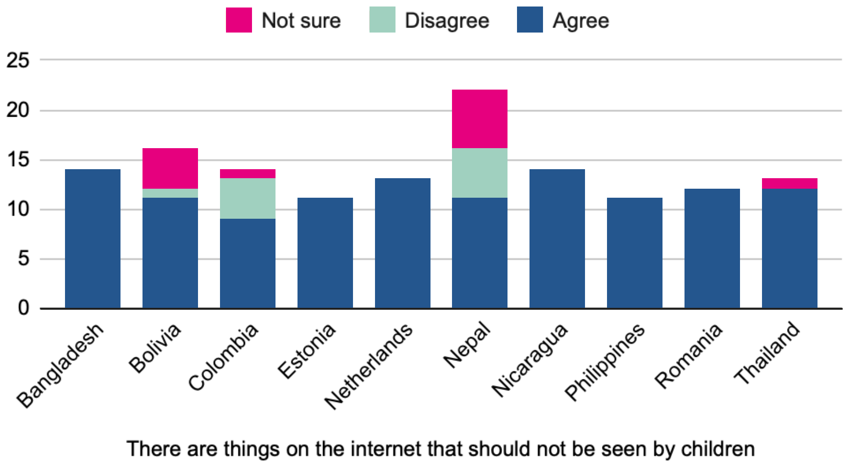
[307] Ghosh, A. K., Badillo-Urquiola, K., Guha, S., LaViola Jr., J. J. and Wisniewski, P. (2018) 'Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control'.

[308] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls.*

of punishment or being further monitored by their parents.[309] It is therefore particularly important to shift away from a surveillance and punishment approach, and instead, develop designs aimed at encouraging and supporting parents in identifying children's real challenges and struggles online and helping them learn how to cope with risks.[310]

According to the latest EU Kids Online survey, 33% of children have seen sexual images online; boys being more likely (37%) compared to girls (29%), while more older children reporting viewing sexual images (61% were 15 to 16 years old, compared to 11% who were 9 to 11 years old).[311] Research indicates that children are as likely to accidentally come across pornography online as they are to intentionally view it.[312] In the focus group discussions, around 84% of participating children indicated that there are things on the internet that should not be seen by children (see Graph 11).

Graph 11. Focus group answers on whether some content is inappropriate for children



Source: Focus groups in 10 countries across Asia, Europe and Latin America

It is crucial to adapt such control to the age and evolving capacities of the child (e.g. a less restrictive approach for older children).[313] A proposed example of a privacy-friendly and respectful parental control approach, potentially suitable for older children, is the Circle of Trust feature, based on the concept of trusted and untrusted contacts lists developed by parents and children. While parents have access to all messages received by their children from untrusted contacts, if the message comes from a trusted sender they can only see content flagged by the app as risky.[314]

[309] Mishna F, Milne E, Cook C, Slane A, Ringrose J. Unsolicited Sexts and Unwanted Requests for Sexts: Reflecting on the Online Sexual Harassment of Youth. Youth & Society. November 2021. https://journals.sagepub.com/doi/10.1177/0044118X211058226.

[310] Wang, G., Zhao, J., Van Kleek, M. and Shadbol, N. (2021) 'Protection or punishment? relating the design space of parental control apps and perceptions about them to support parenting for online safety'.

[311] Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S. and Hasebrink, U. (2020) *EU Kids Online 2020: Survey results from 19 countries*. https://doi.org/10.21953/lse.47fdeqj01ofo.

[312] Belton, E. and Hollis, V. (2016) *A Review of the Research on Children and Young People who Display Harmful Sexual Behaviour Online*. Available at: https://learning.nspcc.org.uk/media/1198/review-children-young-people-harmful-sexual-behaviour-online.pdf (Accessed 25 April 2022).

[313] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls*.

[314] Ghosh, A. K., Hughes, C. E. and Wisniewski, P. (2020) 'Circle of Trust: A New Approach to Mobile Online Safety for Families', 2020 *CHI Conference on Human Factors in Computing Systems.* Honolulu, United States, 25-30 April. ACM, New York, United States, pp.1-14. https://doi.org/10.1145/3313831.3376747.

3.1.3 Parents need help to communicate with their children about risks

Other factors that need to be taken into account when debating parental controls include, among others factors, different models of parenting (e.g. a risk of using overly restrictive control measures at the expense of creating a supportive environment for the child), the **socio-economic background** of the family, specific cultural beliefs or practices, the level of **digital literacy**, and **parental engagement**.[315] A study conducted among parents of young children (aged between 4 and 7 years old) showed that the levels of income and education in the household correlate with the approach towards parental control. For instance, parents with high income and education levels tend to prioritise empowerment and active mediation (i.e. engaging with their children and following up on risky behaviour online), whereas parents with low income and education levels might exhibit more restrictive and controlling strategies while lacking digital media expertise, capacity and confidence to support children in dealing with risks.[316]

In a web-based diary study, Wisniewski, Xu, Rosson, and Carrol (2017) followed 68 teen-parent pairs over the course of two months, looking at online behaviour, risks, and communication about behaviour and risks. They found that **only 28% of teenagers told their parents what they experienced on the internet**. Reasons for not telling their parents were mostly that the teenagers thought that what they had experienced was no big deal or that it would result in an uncomfortable conversation. This finding underlines the notion that teenagers rarely engage with their parents about their life online but rather deal with their experiences on their own.[317] Another survey found that in 40% of instances, children turned to their parents for help after a negative experience online.[318]

> *"Sometime I told my mom that I had a boyfriend and she did not understand me,*
> *so I went to other places where I thought they might want to help me, but those people had other intentions."*
> (Focus group discussions, 15-year-old girl, Bolivia)
>
> *"I play Fortnite and I only talk about that with my friends, my parents don't*
> *understand, there are things you can't talk about with your parents."*
> (Focus group discussions, 15-year-old boy, Bolivia)

3.1.4 Children prefer turning to peers for help

Receiving **peer support** is important for teens. In a recent survey, half of the children surveyed reported having reached out to a friend their own age after a negative experience online.[319] In the focus groups, 56% of participating children also indicated that they are more comfortable talking to friends instead of parents or other adults when encountering something distressing online; 22% disagreed with the statement and 22% were unsure. The children expressed that adults are often more knowledgeable in general, but that telling an adult incurs the risk of punishment or being lectured. Other children indicated that they prefer to speak to friends, as they are more familiar with social media, whereas parents are not.

> *"I feel comfortable sharing with friends, because parents scold us instead of supporting us and they will point out our fault instead of solving our problems. They won't give us mobiles or laptops to use anymore if we share our problem with them."* (Focus group discussions, child)

[315] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls*.

[316] Livingstone, S., Mascheroni, G., Dreier, M., Chaudron, S. and Lagae, K. (2015) *How parents of young children manage digital devices at home: The role of income, education and parental style.* Available at: **http://eprints.lse.ac.uk/63378/1/__lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_EU%20Kids%20Online_EU_Kids_Online_How%20parents%20manage%20digital%20devices_2016.pdf** (Accessed 25 April 2022).

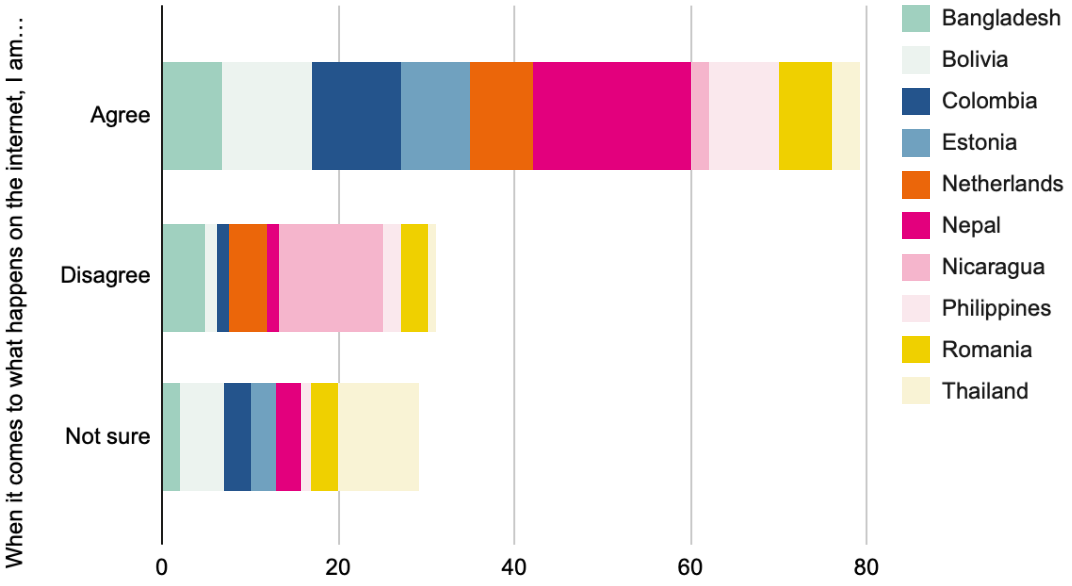
[317] Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2017) 'Parents just don't understand: Why teens don't talk to parents about their online risk experiences', *ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '17), Portland, United States, 25 February – 1 March, pp.523–540. **https://doi.org/10.1145/2998181.2998236**.

[318] Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S. and Hasebrink, U. (2020) *EU Kids Online 2020: Survey results from 19 countries*.

[319] Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S. and Hasebrink, U. (2020) *EU Kids Online 2020: Survey results from 19 countries*.

Children participating in the focus groups also showed a preference for turning to friends for advice when facing risks online. Most of the participants agreed with the statement *"When it comes to what happens on the internet, I am more comfortable talking to my friends instead of my parents or adults"*.

Graph 12. Focus groups answers on seeking advice when facing risks online



Source: Focus groups in 10 countries across Asia, Europe and Latin America

Children in the Philippines expressed that they think their friends can more easily relate to them because they share the same experience, age, mindset, and feelings. Those thoughts were echoed across other countries. Children worded it as follows:

> *"It is not easy to talk, sometimes parents make the mistake of scolding, and so I don't want to tell them. That is why sometimes we look for people to listen to us and understand us."*
> (Focus group discussion, 13-year-old girl, Bolivia)
>
> *"Some adults are not very understanding or they don't know much about technology, even more so when compared to friends, who can understand better."*
> (Focus group discussion, 13-year-old boy, Romania)

Where a child fears possible backlash from friends after something happened, this support could also be sought in an anonymised environment such as a moderated social **support platform**.[320]

[320] Hartikainen, H., Razi, A. and Wisniewski, P. (2021) 'Safe Sexting: The Advice and Support Adolescents Receive from Peers Regarding Online Sexual Risks'. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 42, pp.1-31 **https://doi.org/10.1145/3449116**.

3.1.5 Online risks also come from peers

Adolescence is a time for exploring sexuality. Online platforms provide an outlet for such exploration. While it can be healthy and safe to engage in sexual activities online, these present risks of online harm and can foster online violence through sexual extortion, non-consensual sexting, and grooming. Teenagers receive sexual solicitation from **both peers and strangers**. In a 2022 global survey, 31% of girls and 12% of boys had received sexually explicit content from a stranger. Similarly, 24% of girls and 12% of boys had an unknown person asking them to do something sexually explicit online that they were uncomfortable with or that they did not want to do.[321] When it came to sexual interactions with peers, 33% of respondents to the global survey had received sexually explicit content from peers (37% of boys and 25% of girls). Boys were also more likely to find the experience of receiving sexual content from peers as a positive experience, while girls were, for the most part, uncomfortable with it.[322] A key difference between receiving content from strangers or from people known to the child is that it is **easier for the child to reject and report solicitations from strangers** than from people they know.[323]

3.1.6 Online spaces may feel private to children leading to over-disclosure

As mentioned above, adolescence is a time of risk-taking and children often underestimate the risks they face in disclosing information and sharing content. Research suggests that teenagers are not so concerned about privacy and are *"less engaged in privacy management"*.[324] Teenagers, while aware of privacy risks, value the social and relational benefits of disclosing information and publishing content online over the potential risks. This is due to their need for social interaction.[325] Boys in particular are less concerned about privacy than girls, which may be linked to the heightened risks of online harms girls face, parental intervention, as well as socialisation. Studies show that girls were more likely to have had a discussion about online privacy with their parents than boys.[326]

Source: Focus group in Estonia

[321] Economist Impact and WeProtect Global Alliance (2022) ***Estimates of childhood exposure to online sexual harms and their risk factors.*** Available at: **https://www.weprotect.org/economist-impact-global-survey/#report** (Accessed 18 April 2022).

[322] ibid

[323] Hartikainen, H., Razi, A. and Wisniewski, P. (2021) 'Safe Sexting: The Advice and Support Adolescents Receive from Peers Regarding Online Sexual Risks'.

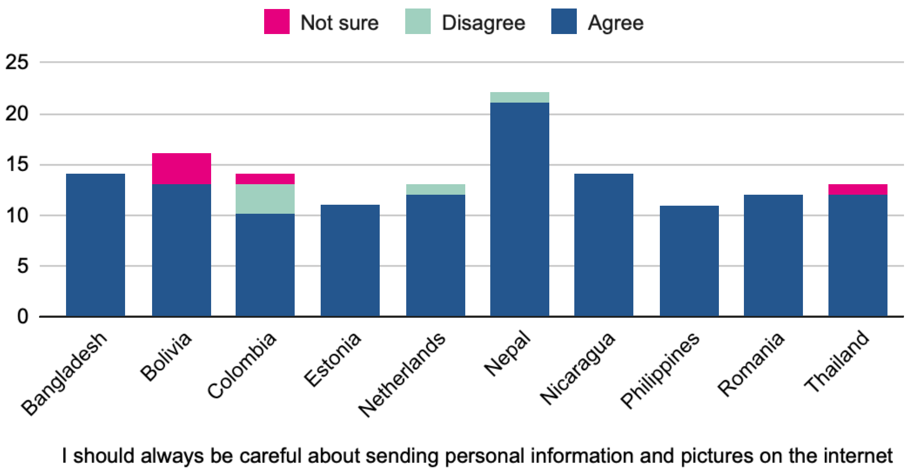
[324] Stoilova, M., Livingstone, S. and Khazbak, R. (2021) *Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children's internet use and outcomes*. Available at: **https://www.unicef-irc.org/publications/pdf/Investigating-Risks-and-Opportunities-for-Children-in-a-Digital-World.pdf** (Accessed 27 April 2022).

[325] Pinter, A. T., Wisniewski, P., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Adolescent Online Safety: Moving Beyond Formative Evaluations to Designing Solutions for the Future'.

[326] ibid

In the conducted focus group discussions, the children displayed the ability to recognise the risks of sending out personal information or photos, and agreed that they should be careful in doing so. In all countries, a clear far majority agreed with the given statement. Only in Colombia, a larger proportion, but still the minority of the group disagreed that they should always be careful about sending personal information and pictures on the internet.

Graph 13. Focus group answers on sending personal information or content



Source: Focus groups in 10 countries across Asia, Europe and Latin America

Livingstone et al. (2019) makes the distinction between three privacy contexts: 1. **interpersonal privacy**; 2) institutional privacy; and 3) commercial privacy. Interpersonal privacy is abouts relationships between individuals as well as groups. It relates to privacy decision and practices in the online environment, which is the most relevant to online safety.[327] Children's privacy decisions online are influenced by gender, parents, peers, their interpretation of the social situation, their attitude towards privacy, prior negative experiences, their social media use, their **digital literacy** in navigating privacy features, as well as the design of the online environment.[328] In general, children are more likely to share personal information and feelings online. Privacy breaches do not necessarily lead to a stress response as children are also more used to their data being provided to third parties and take fewer precautions to protect their privacy.[329]

Children's experience of online environments does not necessarily imply the same privacy considerations as adults. In fact, the use of online public spaces may be perceived as offering some privacy to children as they are parent-free. On platforms where the child only interacts with peers, the child may experience the platform as a private space and be inclined to share more content. Some online spaces may also feel more private if the child has fewer known contacts on that platform and they are less under the influence of family- or peer-related social norms. Some online spaces also offer the ability to communicate in a way that carries some closeness and trust, such as anonymous peer support platforms or private blogs.[330] This varied experience of the private sphere may impact children's approach to privacy and vulnerability to risks.[331] The visibility of children's activities and actions within certain online spaces affects their perception of privacy and their

[327] Livingstone, S. Stoilova, M. and Nandagiri, R. (2019) *Children's data and privacy online: Growing up in a digital age*. An evidence review. Available at: https://www.lse.ac.uk/media-and-communications/assets/documents/research/projects/childrens-privacy-online/Evidence-review.pdf (Accessed 29 April 2022).

[328] ibid

[329] Stoilova, M., Livingstone, S., and Khazbak, R. (2021) 'Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children's internet use and outcomes'.

[330] Livingstone, S. Stoilova, M. and Nandagiri, R. (2019) *Children's data and privacy online: Growing up in a digital age. An evidence review*. Available at: https://www.lse.ac.uk/media-and-communications/assets/documents/research/projects/childrens-privacy-online/Evidence-review.pdf (Accessed 29 April 2022).

[331] ibid

privacy decisions. In turn, the online audience uses the visible information to form impressions and monitor children, as well as to seek sexual interactions with them. Children are still learning how to navigate appropriate information disclosure online and therefore need extra protection and support.[332]

> *"When I opened my Facebook account, I didn't understand how to use and chat with others. A boy wanted my picture and I send my hand picture. Later I understood that this is very harmful for me."*
> (Focus group discussions, 17-year-old girl, Bangladesh)

3.2 INDUSTRY SOLUTIONS THAT WORK FOR CHILDREN

3.2.1 Evidence-based solutions

Drawing on the literature, the two senior experts involved in the current research identified the child safety designs best suited to protect children specifically from OSEC (i.e. grooming and sextortion) and CSAM, including self-generated CSAM. The senior experts evaluated the identified child safety by design measures against a set of criteria:

1. Effective in preventing OSEC/CSAM, reducing risks of inappropriate contacts, and reducing inappropriate exchange of images, videos or chats.
2. Unlikely to be circumvented by children.
3. Do not infringe on user privacy and minimise privacy risks.
4. Are in line with children's rights (e.g. right to be heard, right to participate in society and in decisions affecting them, best interests of the child).
5. Do not prevent children from accessing digital environment in a safe way.

In addition, eight external senior experts provided written inputs into the identified solutions. The participants in the online workshop also discussed the solutions aimed at the industry. This feedback has been included in this Section. One overall point they raised is that **no single solution can sufficiently or effectively tackle OSEC**. Each solution must be part of a suite of solutions aiming to holistically tackle online harms from various angles.

The following industry solutions are discussed in detail below:

1. Deploy parental control features adapted for children across all age groups
2. Develop features that encourage child or teen empowerment and self-regulation
3. Deploy intelligent privacy features customised to keep all children safe
4. Strengthen age verification and assurance by using multiple sources
5. Implement technologies that identify risks combined with risk mitigation strategies

The industry solutions are focused on technology-based interventions. Technology can be used to tackle OSEC through various types of interventions, each having its own merits and drawbacks. These can consist of individual apps, design techniques, platform features, or AI tools. This section looks at technology interventions in terms of the safety approach they can provide.

1. DEPLOY PARENTAL CONTROL FEATURES ADAPTED FOR CHILDREN ACROSS ALL AGE GROUPS

RECOMMENDATION 1: The industry should deploy parental control and safety features adapted to the various age groups of their users. It should include more restrictive parental control features for younger children, and category-based parental control features that enhance active mediation for teenagers.

SOLUTION IN DETAIL

- **Parental monitoring and restriction for children aged 3 to 12 years old, including lightweight platform-native parental oversight and parental control apps.** Parental oversight and monitoring should be included at a light level on platforms or through third-party apps for younger children as it may be too intrusive for older children. Such parental monitoring and restrictions would help keep younger children from falling

[332] Ghosh, A. K., Badillo-Urquiola, K., Guha, S., LaViola Jr., J. J. and Wisniewski, P. (2018) 'Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control'.

victim to harmful situations and give them immediate support when needed. A potential drawback is that many parental monitoring and restrictions are too heavy-handed and do not preserve privacy, therefore lightweight or category-based versions are recommended (see below).

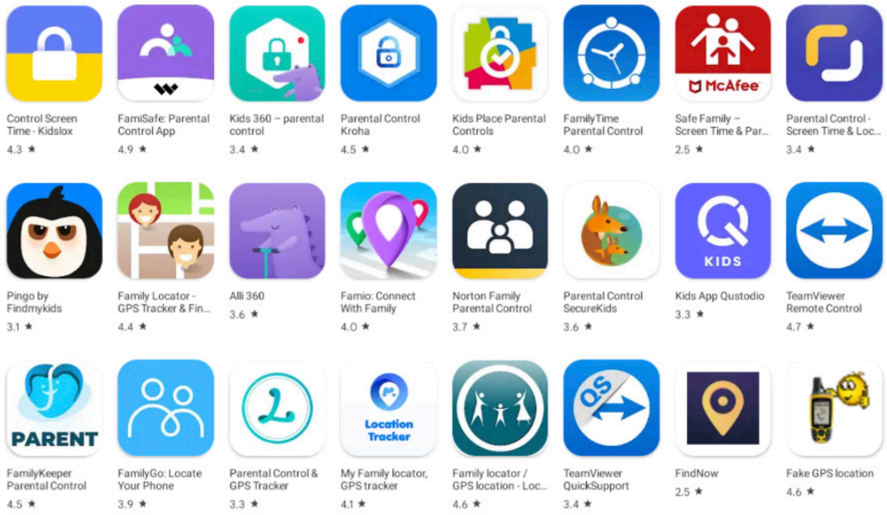
- **Category-based parental controls for all children.** Parental control apps should provide lightweight or category-based risk detection to keep children safe, instead of monitoring the entirety of children's online activities. Such apps offer privacy-preserving protection by providing meta-level information of children's online activities rather than a full disclosure of their online use and the content of their interactions. For instance, the parental control app would only provide high-level summaries of children's activities grouped by app category, such as apps used, top contacts, and time spent on the device per day. Category-based filters should not offer blanket restrictions, due to the risk of over-blocking access. Such filters should be accompanied by active mediation tools and should avoid restrictive approaches. Category-based apps used without active mediation could give parents a false sense of security that their children are protected, reducing opportunities for parental communication as well as limiting children's opportunities.
- **Parental control supporting active mediation for children aged 13 to 17 years old, recognising teenagers needs for more autonomy.** Apps should be designed to support active parental mediation by enhancing parent-children communication around technology use and its limits, and involving parents in supporting children facing potential online risks and harms, coupled with engaged supervision and flexibility. Parental control and safety features should be designed with teenagers as their end users, rather than being parent-centric. This would help increase teenagers' adoption of safety features or apps. They should include functionalities beneficial to teenagers and that help them to negotiate with their parents.

RATIONALE

The above solutions have the potential to meet parents' needs to ensure the safety of their children online, while meeting teenagers' needs for privacy through the use of apps that employ a level of abstraction in the information they provide about children's activities online.

Parental control is often mentioned as a safety tool to protect children against online harm, including OSEC. Usually involving apps downloaded by both parents and children, parental control can take various forms, from blocking and filtering certain apps or websites to limiting screen time or monitoring children's online activities and interactions. Parental control apps have the ability to limit the time children spend online and the risks of being exposed to harmful content. However, research shows mixed results in terms of the effectiveness of parental control apps, either because children are able to circumvent the app or because these apps are overly restrictive and harm the trust between parents and children.[333]

Figure 3. Parental control apps in Google Play


Control Screen Time - Kidslox
4.3 ★
FamiSafe: Parental Control App
4.9 ★
Kids 360 – parental control
3.4 ★
Parental Control Kroha
4.5 ★
Kids Place Parental Controls
4.0 ★
FamilyTime Parental Control
4.0 ★
McAfee
Safe Family – Screen Time & Par...
2.5 ★
Parental Control - Screen Time & Loc...
3.4 ★
Pingo by Findmykids
3.1 ★
Family Locator - GPS Tracker & Fin...
4.4 ★
Alli 360
3.6 ★
Famio: Connect With Family
4.0 ★
Norton Family Parental Control
3.7 ★
Parental Control SecureKids
3.6 ★
KIDS
Kids App Qustodio
3.3 ★
TeamViewer Remote Control
4.7 ★
PARENT
FamilyKeeper Parental Control
4.5 ★
FamilyGo: Locate Your Phone
3.9 ★
Parental Control & GPS Tracker
3.3 ★
Location Tracker
My Family locator, GPS tracker
4.1 ★
Family locator / GPS location - Loc...
4.6 ★
QS
TeamViewer QuickSupport
3.4 ★
FindNow
2.5 ★
Fake GPS location
4.6 ★

Source: Screenshot of a selection of parental control apps in Google Play

[333] Badillo-Urquiola, K., Chouhan, C., Chancellor, S, De Choudhary, M. and Wisniewski, P. (2019) 'Beyond Parental Control: Designing Adolescent Online Safety Apps Using Value Sensitive Design', ***Journal of Adolescent Research***, 35(1), pp.147-175. https://doi.org/10.1177/0743558419884692; Wang, G., Zhao, J., Van Kleek, M. and Shadbol, N. (2021) 'Protection or punishment? relating the design space of parental control apps and perceptions about them to support parenting for online safety'.

The use of technologies often leads to the **tensions between parents and children**. The literature reports that parents tend to underestimate their children's social media use, while children felt that parents did not guarantee their right to privacy. With an approach to social media app usage that is deemed too restrictive, family rules around it are often broken by parents and children.[334] Previous research has shown that the majority of safety features support parental control (89%) rather than self-regulation (11%), while teenagers (79%) tend to dislike overly restrictive parental control features.[335]

While restrictive parental control apps may reassure parents, they are therefore not the ideal solution for protecting children.[336] As Phippen and Brennan (2021) put it, by *"imposing an unacceptable level of control and data collection upon them [..] what we are actually wishing to achieve is not safe children, but compliant ones"*.[337]

Engaging in **co-use, or co-viewing, and active parental mediation** is a protective factor against online risk. Active parental mediation is positively associated with online opportunities and **lower risks**, while restrictive parental mediation is associated with higher online risks and fewer online opportunities, especially with teenagers.[338] Active parental mediation is a technique where parents talk about their children's technology use and the associated risks with them, and actively engage them in discussions. Co-use/co-viewing mediation refers to members of the same household using or watching the technology at the same time.

Trust becomes an essential component of the relationship between parents and teenagers.[339] Trust does not imply an absence of monitoring and information disclosure, but requires a more cooperative type of parental mediation that respects teenagers' needs for privacy and autonomy. The literature cautions against overprotecting teenagers and paternalist approaches, which may hinder their capacity to learn to protect themselves. Approaches that support **parent-teenager communication** and **trust building** are shown to be more effective than control-based approaches.[340] Trust in parent-child relationships contributes as a protective factor to online harms as children will more easily turn to their parents for support when facing risks. Research indicates that parents who were active in their teenagers' online activities and trusted that they behaved reasonably online have a more accurate insight into their children's online experiences.[341]

Therefore, parental control apps that support active and/or co-use parental mediation appear more effective in reducing online risks, while parental control apps supporting restrictive mediation are negatively associated with online risks, especially for teenagers. Parental control apps would need to account for regional and/or cultural differences (i.e. more restrictive and more liberal societies). These apps also require a level of parental digital literacy to effectively manage such processes. Therefore, they must also provide easy-to-use interfaces for parents, prompting them towards effective ways to engage with their children regarding online risks.[342]

Parental control apps supporting active and co-use parental mediation can be an effective tool to reduce risks of exposure to OSEC. A more restrictive parental control may however be appropriate and effective for younger children (under 13 years old).

A last consideration for designers is to ensure that safety features cover **all types of risks, including from peers and family members**. Considering that sexual solicitation from people children know, including peers, is a common occurrence, features that are overly geared towards stranger danger situations do not adequately support children when they encounter online harms from people they know.

[334] Ghosh, A. K., Badillo-Urquiola, K., Guha, S., LaViola Jr., J. J. and Wisniewski, P. (2018) 'Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control'.

[335] ibid

[336] ibid

[337] Phippen, A. and Brennan, M. (2021) ***Child Protection and Safeguarding Technologies Appropriate or Excessive 'Solutions' to Social Problems?*** Routledge, New York.

[338] Cabello-Hutt, T., Cabello, P. and Claro, M. (2018) 'Online opportunities and risks for children and adolescents: The role of digital skills, age, gender and parental mediation in Brazil'.

[339] Badillo-Urquiola, K., Chouhan, C., Chancellor, S, De Choudhary, M. and Wisniewski, P. (2019) 'Beyond Parental Control: Designing Adolescent Online Safety Apps Using Value Sensitive Design'.

[340] Pinter, A. T., Wisniewski, P., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Adolescent Online Safety: Moving Beyond Formative Evaluations to Designing Solutions for the Future'.

[341] Erickson, L. B., Wisniewski, P., Xu, H., Carroll, J. M., Rosson, M. B. and Perkins, D. F. (2016) 'The boundaries between: Parental involvement in a teen's online world'.

[342] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

2. DEVELOP FEATURES THAT ENCOURAGE CHILDREN'S EMPOWERMENT AND SELF-REGULATION

RECOMMENDATION 2: The platforms and industry designers should develop and deploy safety features that build on children's needs to learn to identify and deal with risks, and that empower them to self-regulate their own online behaviour.

SOLUTION IN DETAIL

- **Self-regulation and self-empowerment (especially for children aged 13 to 17 years old).** Safety features should recognise teenagers' needs for autonomy, privacy, and agency by empowering them to protect themselves when using technologies. The features should include designs that help teenagers self-regulate their own behaviour online and manage online risks. By taking a more teenager-centric approach to online safety, designers can foster teenagers' sense of personal agency in dealing online risks, helping them learn vital skills for safe online engagement.[343] Design features could include a parent/teenager interface with parents and children each having their own account and access, thus promoting shared ownership over safety with the ability to customise features according to the identified needs. Through the interface, parents and children could assign risk ratings, or flag risks or contacts that may be unsafe. Prompts could help parents and children establish communication about, and an approach to online safety.[344]
- **Peer support, including online peer support platforms.** Providing children a safe place to get peer and/or professional support away from general (public) forums can be an effective way to help teenagers learn and empower themselves. Online peer support platforms should have child-centric designs and ensure a safe space for children, including vulnerable young people. Peer support platforms should require proof of identity from members while supporting anonymity during the interactions. Such platforms must also be moderated by trained professionals and supported with technology and safety enhancing designs (e.g. voting up or down advice, educational features and reporting mechanisms).

RATIONALE

Studies show that children understand the need for safety and parental oversight. Children report needing support from trusted adults to deal with online risks. At the same time, children, especially teenagers, also need personal agency and privacy when using social media apps.[345]

Teenagers tend to underestimate risk and their ability to avoid risk online. Generally, they are more likely than adults to engage in **risk-taking behaviour**. They have a growing need for autonomy, privacy and self-regulation.[346] Cognitively, teenagers are developing their ego identity as individuals separate from the family unit. This leads to a need for independence and seeking new social experiences. Risk-taking behaviour is a part of this process. As a result, the parent-child relationship evolves in adolescence from unilateral control to **cooperation and negotiation**.

Solutions that rely on teenagers' empowerment and agency are thus more appropriate to their stage of cognitive development. The literature suggests that these types of solutions are more effective in protecting them from online risks. Models focusing on supporting teenagers in learning from their own risk-taking behaviours, **self-monitoring, impulse control** and developing **risk-coping strategies** provide an avenue to addressing teenagers' risk-taking behaviour and helping them to protect themselves against online risks through developing their **self-regulatory** competence.[347] Studies with children indicate that they prefer a self-regulatory approach to assisting them in dealing with online risks (e.g. self-monitoring features). Children as young as 7 years old have shown awareness of online risks, such as the use of fake accounts for malicious purposes, and at the same

[343] Ghosh, A. K., Badillo-Urquiola, K., Guha, S., LaViola Jr., J. J. and Wisniewski, P. (2018) 'Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control'.

[344] Badillo-Urquiola, K., Chouhan, C., Chancellor, S, De Choudhary, M. and Wisniewski, P. (2019) 'Beyond Parental Control: Designing Adolescent Online Safety Apps Using Value Sensitive Design'.

[345] Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewski, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online', *18th ACM International Conference on Interaction Design and Children*, Boise, United States, 12-15 June. ACM, New York, United States, pp.394-406. https://doi.org/10.1145/3311927.3323133.

[346] Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors', *ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '15), Vancouver, Canada, 14-18 March. ACM, New York, United States, pp.583-599. https://doi.org/10.1145/2675133.2675287.

[347] Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors'; Ghosh, A. K., Badillo-Urquiola, K., Guha, S., LaViola Jr., J. J. and Wisniewski, P. (2018) 'Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control'.

time, know that they need support on how to deal with online risks. Self-regulatory features can take the form of personal privacy features, asking for help from parents/trusted adults, or intelligent assistance.[348] Such a self-regulatory approach is beneficial for teenagers' development in helping them identify potential and hidden risks and encouraging them to seek support when needed.[349]

A self-regulatory/empowerment approach is grounded in the understanding that it is not realistic to shield teenagers from online harms, due to the pervasiveness of online risks and teenagers' needs for autonomy. Solutions must empower them to address the risks they may encounter and to become resilient
against them.[350]

Co-designing self-regulatory and empowering features together with children of various age groups and vulnerabilities can ensure that the features account for the needs of a wide range of children. Not all children have the same maturity, capacity, and resilience to self-protect. Therefore, such an approach must be combined with other protective measures as children should not ultimately be responsible for protecting themselves from harm.[351]

The design features could enable children to **identify** and **manage** low-level risks by themselves, while making it easier to reach out to external help systems (parents, support services, reporting mechanisms) for high-level risks.[352] For example, a social media app could include a feature where **children can notify** a parent about a potentially risky situation they encounter. Similarly, social media apps could include online safety resources, such as educational videos that teach children to report OSEC or how to change privacy filters, to guide parents and children through a dialogue on safety and an approach of **shared responsibility**.[353]

Self-regulatory/empowerment approaches also help teenagers to seek support in ways that suit their needs. When it comes to sexuality and sexual exploitation or abuse, teenagers may be reluctant to seek the support of a parent or adult due to the shame about the risk they took or encountered and fear of judgment. **Peer online support** can then become an important resource for teenagers. This is especially true for vulnerable teenagers, such as LGBTQI+ teens and survivors of sexual abuse. The literature reports that vulnerable teenagers are more likely to seek support from peers through online platforms and forums.[354]

Through peer support, teenagers are empowered to discuss the online risks or harms they face with peers who have also encountered similar situations, building a sense of community.[355] Studies show that peers provide information, offer advice on how to handle or mitigate the sexual abuse/risk, and give emotional support.[356] The **anonymity** of the platforms allows them to discuss **sensitive issues**, yet it can also pose some risks of bullying, victim-blaming or retaliating in what is meant to be a safe space. Therefore, peer support platforms should be designed with the safety of children in mind. This could entail **restricting access** by age and affinity (e.g. LGBTQI+). Peer support platforms could also require proof of identity to access them, while supporting anonymity during the interactions. In addition, such safe spaces should be **moderated** by trained professionals who are supported with technology and human-centred algorithms.[357] Features such as the ability for **users to**

[348] Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewski, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online'.

[349] Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors'.

[350] McHugh, B. C., Wisniewski, P., Rosson, M. B. and Carroll, J. M. (2018) 'When social media traumatizes teens: The roles of online risk exposure, coping, and post-traumatic stress', *Internet Research*, 28(5), pp.1169-1188. https://doi.org/10.1108/IntR-02-2017-0077.

[351] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[352] Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors'.

[353] Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewski, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online'; McHugh, B. C., Wisniewski, P., Rosson, M. B. and Carroll, J. M. (2018) 'When social media traumatizes teens: The roles of online risk exposure, coping, and post-traumatic stress'.

[354] Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2020) 'Let's Talk about Sext: How Adolescents Seek Support and Advice about Their Online Sexual Experiences', *2020 CHI Conference on Human Factors in Computing Systems*, Honolulu, United States, 25-30 April. ACM, New York, United States. https://doi.org/10.1145/3313831.3376400.

[355] Hartikainen, H., Razi, A. and Wisniewski, P. (2021) 'Safe Sexting: The Advice and Support Adolescents Receive from Peers] Regarding Online Sexual Risks'.

[356] ibid

[357] Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2020) 'Let's Talk about Sext: How Adolescents Seek Support and Advice about Their Online Sexual Experiences'.

vote up or down advice provided by peers can offer a measure of trustworthiness, with the peers considered most helpful being rewarded with a higher ranking. This can be particularly effective as collective judgements of social groups tend to be more accurate than individual judgements.[358]

Technology could also support such safe spaces, through development of virtual companions and AI therapists, which already have a significant number of users today, and are rapidly becoming more common as technology progresses. These technologies can create a low threshold for children to reach out for help on their issues, and be a stepping stone to talking with a real therapist.[359]

3. DEPLOY INTELLIGENT PRIVACY FEATURES CUSTOMISED TO KEEP ALL CHILDREN SAFE

RECOMMENDATION 3: Platforms should deploy intelligent privacy features to ensure the highest privacy settings for any child and make settings customisable for certain age groups. Those features should include intelligent privacy settings by default for all user accounts under 18 years of age, with proactive privacy features as well as reactive privacy features.

SOLUTION IN DETAIL

- **Intelligent privacy by default.** The status quo bias results in most users, including children, and in particular younger children, opting to leave default privacy settings. Platforms should adapt default privacy settings for children's accounts, using intelligent settings. This can take the form of defaulting a child's profile to 'friends only' rather than 'public'. Platforms, such as Instagram, have recently made the decision to not allow strangers who do not follow teenagers on Instagram to send them private messages. This could be part of intelligent defaults instead of an absolute restriction, as such restrictions may unintentionally harm vulnerable young people, such as those in the LGBTQI+ community who have different social support needs.
- **Proactive privacy features.** Such features would nudge child users to proactively manage their privacy. This can take the form of prompts to regularly review privacy settings. The use of proactive privacy features, such as sensitivity filters and other computational approaches to detect inappropriate content can also help prevent unintentional exposure to sexually explicit content and CSAM. These features should empower young people to make protective (i.e. opt-in) decisions, rather than making them feel like they are being censored in online spaces. Censoring approaches may lead children to seek content in riskier online spaces, such as the dark web.
- **Retroactive privacy features.** These include technologies that allow users to report problems (including to authorities) after negative behaviour occurs, blocking users, or getting assistance. It is crucial that children are given the ability to make and correct privacy-related mistakes online. Such technologies support the empowerment of children to respond to online risk, providing them with the agency to take action and ask for assistance. As such, platforms should provide young people with retroactive privacy features, such as the ability to untag, delete, block, and report inappropriate content and people once their privacy or safety concern has become heightened.

RATIONALE

Privacy can be a safety tool against online risks, including grooming and sexual extortion. Stricter privacy settings, including limiting disclosure of information to a selected contact group, are appropriate for young children, in order to minimise interactions they may have with unknown people. Research shows that younger children (under 12 years of age) have more difficulties managing their online privacy settings than children from 12 years and above, and find it challenging to understand abstract notions such as 'privacy' and 'safety'. This is especially the case for children under eight years old.[361] Stricter privacy approaches are not sufficient to keep young children safe. They must also be accompanied by other safety tools, as we know that younger children are more at risk from people they know.

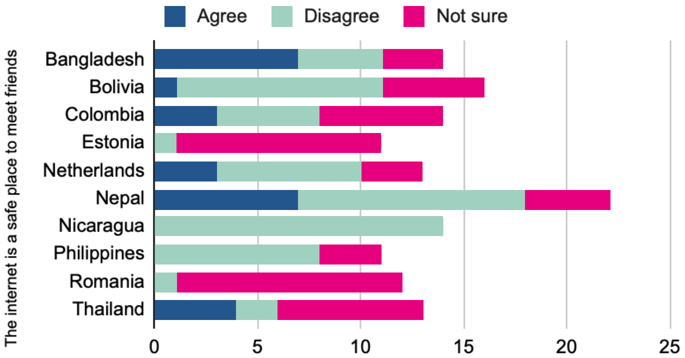
[358] Hartikainen, H., Razi, A. and Wisniewski, P. (2021) 'Safe Sexting: The Advice and Support Adolescents Receive from Peers Regarding Online Sexual Risks'.

[359] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[361] Livingstone, S. Stoilova, M. and Nandagiri, R. (2019) ***Children's data and privacy online: Growing up in a digital age. An evidence review***.

As mentioned above, children's experience of privacy online may vary depending on the platforms they are active on, where parent-free spaces feel more private. Teenagers, in particular, are prone to over-disclosing information, prioritising the social benefits over privacy. Sharing personal content and information online, having a public profile, including unknown people as social media contacts, and chatting with unknown people constitute online risk behaviours associated with online grooming and sexual exploitation.[362] In the focus group discussions, children were asked if they thought the internet was a safe place to meet new friends. Only in Bangladesh, the majority of children felt it was safe. In Bolivia, the Netherlands, Nepal, Nicaragua, and the Philippines, the majority of children felt it was not safe to meet new friends online. Children thus indicate that there are certain privacy risks online. Design features that address those privacy risks can help reduce risks of exposure to OSEC.

Graph 15. Focus group answers on safety of making new friends online



Source: Focus groups in 10 countries across Asia, Europe and Latin America

Design features can also be a source of risks. For example, the ability to tag other users without prior consent and location-based features, such as automated geotagging and GPS-enabled tracking, facilitate privacy breaches.[363] The ease of creating fake social media accounts facilitate grooming and monitoring of children's activities. There is also the risk that data collected from children's activities is sold to various agents (e.g. advertisers) and data breaches can lead to children's personal data being exposed to perpetrators.[364] In addition, the reliance on self-reporting of sexually explicit content means that content removal is often delayed and inconsistent across platforms, which facilitates the dissemination of the content and the risks of exposure.[365] As we know, exposure to sexually explicit material and CSAM affects children and influences their own exploration of sexuality, which may lead them to seek such material themselves.

Research shows that social media users tend to rarely change the **default privacy settings**. The reason for this behaviour lies in a bias for the status-quo.[366] Where additional privacy options are available (e.g. deciding for each post who can see it), most users tend to stick with their defined default settings.[367] Studies indicate that users are easily influenced by the platform's design in the type of information and amount of content they disclose, including nudging towards increased disclosures and content sharing. Default settings can lead to over-disclosure and can be interpreted as recommended settings. The ability to control privacy settings at granular level (e.g. for each post) can give a greater sense of control and thus a lesser concern over privacy,

[362] Stoilova, M., Livingstone, S. and Khazbak, R. (2021) *Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children's internet use and outcomes.*

[363] McHugh, B. C., Wisniewski, P., Rosson, M. B. and Carroll, J. M. (2018) 'When social media traumatizes teens: The roles of online risk exposure, coping, and post-traumatic stress'.

[364] Livingstone, S. Stoilova, M. and Nandagiri, R. (2019) ***Children's data and privacy online: Growing up in a digital age. An evidence review.***

[365] ibid

[366] Joeckel, S., Dogruel, L. (2019) 'Default effects in app selection: German adolescents' tendency to adhere to privacy or social relatedness features in smartphone apps', *Mobile Media & Communication*, 8(1), pp.22-41. **https://doi.org/10.1177/2050157918819616** (Accessed 8 May 2022).

[367] Fiesler, C., Dye, M., Feuston, J. L, Hiruncharoenvate, C., Hutto,C. J., Morrison, S., Khanipour Roshan, P., Pavalanathan, U., Bruckman, A. S., De Choudhury, M. and Gilbert, E. (2017) 'What (or Who) Is Public? Privacy Settings and Social Media Content Sharing', *ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '17), Portland, United States, 25 February – 1 March. ACM, New York, United States, pp.567–580. **https://doi.org/10.1145/2998181.2998223**

leading to over-disclosure of information.[368] This also applies to children's behaviour online. A study of German adolescents showed that they tend to adhere to the default settings. Social media platforms in general tend to be designed with disclosure of information and content as a default, with some privacy options. When adolescents are offered the opposite approach, i.e. privacy by default and disclosure as optional, they similarly follow the default settings, as they do on self-disclosure by default platforms.[369]

Gender and age also affect the tendency to opt for more privacy, with girls more likely to opt for privacy options than boys, while younger children are more likely to follow the default settings.[370] Research shows that younger children have knowledge gaps when it comes to risk and privacy and are thus less apt to proactively manage their privacy.[371] Default privacy settings set to high privacy levels are helpful for younger children who do not fully understand how the settings work, and for vulnerable children whose parents are not able to guide them in making changes to settings. Automatically applying high privacy settings also removes the burden of safety from children and their parents, shifting the onus of responsibility for designing safe spaces to platforms instead of individuals.[372]

Concern about privacy does not trigger less disclosure of information online, also called the 'privacy paradox'. Exposure to risk appears to be more effective in changing privacy behaviour. This explains the higher likelihood of girls to be concerned with privacy and to engage in proactive privacy management, as they are more exposed to risks and harm.

Proactive privacy management (e.g. regularly reviewing privacy settings, being careful about information disclosed and content posted, untagging pictures, being careful about who to friend, setting friends only view/access, etc.) and better privacy defaults are key to minimising exposure to risks.[373] Proactive privacy management can help prevent risks before they materialise by ensuring that adults are not able to interact with children online inappropriately. This can be achieved through regularly prompting child users to review their privacy settings or when they are about to post publicly to warn about risks.

Retroactive privacy features are also a vital set of solutions that empower children to manage risks and react to harm. These include safety features such as blocking, muting, and reporting. This provides children with an appropriate level of control in dealing with risks. Platforms should deploy such features as industry standard and ensure adequate and speedy follow-up to reporting, especially when a child user files a report.

4. STRENGTHEN AGE VERIFICATION AND ASSURANCE BY USING MULTIPLE SOURCES

RECOMMENDATION 4: Platforms should strengthen their age verification systems using multiple sources of information and data. In addition, the industry should define standards for age verification mechanisms and risk-driven age verification implementations based on best practices.

SOLUTION IN DETAIL

- **Strengthen age verification systems during registration.** Age verification is mostly done through easy-to-circumvent self-declaration and later checked through certain measures. This leads to many children being active on platforms that are not intended to be used by them. Age verification during the registration phase
  - on platforms should be stronger to prevent children from access such platforms by using:
  - officially provided data, such as passports, visas, or medical records;

    public databases.

[368] Acquisti, A., Brandimarte, L. and Loewenstein, G. (2022)' Privacy and Behavioral Economics'. In: Knijnenburg, B.P., Page, X., Wisniewski, P., Lipford, H.R., Proferes, N., Romano, J. (eds) Modern Socio-Technical Perspectives on Privacy. Springer, Cham. https://doi.org/10.1007/978-3-030-82786-1_4.

[369] Joeckel, S., Dogruel, L. (2019) 'Default effects in app selection: German adolescents' tendency to adhere to privacy or social relatedness features in smartphone apps'.

[370] ibid

[371] Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewski, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online'.

[372] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[373] Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors'.

- **Strengthen age verification checks after registration.** After the initial age verification check, platforms should increase their efforts to check whether the age provided during the registration phase is correct by using:
  - automatically generated data, such as haptics, motion analysis and user habits;
  - user-reported data, such as data provided by child users, their parents, or other users of the app to verify their identity.

RATIONALE

As explained in the Section on the EU framework, the Audio-visual Media Services Directive solely requires **age verification** to be in place, but does not require it to be effective. Neither are there currently any national laws that provide guidelines about what effective age verification is. Most social media platforms use an age limit of 13 years old, where they ask the new user for a **self-declaration** of their age. This measure is easy to circumvent, as children under the age of 13 can lie about their age in order to gain access to the platform. As a consequence, these children are exposed to risks and content that is not suitable for their age.

> *"It is right to lie about age online, because young people cannot access Facebook, YouTube, TikTok and online gaming."* (Focus group discussion, child, Nepal)

Age verification methods can be used to protect children from a variety of online risks. Lessons learnt from methods used to protect children in the context of online gambling, gaming and shopping can be valuable in strengthening these measures during the phase of registering on a platform. The first is the use of **ID verification led by the government**. An official service could be provided by the government where government databases are checked to verify identities. This is done in Denmark, Italy, and Spain for instance. Using official documents, such as passports, visas, or medical records could also contribute to age and identity verification. Additionally, there are companies that check public databases to see if someone is listed. Sources say that 85 to 90% of adults are recorded in a database somewhere. No data is available for children, but the expectation is that this number will be lower.[374] It is important to note that the use of ID verification would create significant obstacles in countries outside the global North. In addition, in some countries, this approach can create risks of authoritarian abuse and restrictions on children's access to the digital environment. It could also lead to inequality of access, leaving behind the most vulnerable children with higher risks for OSEC, such as refugee children and those children who might not want to use their officially provided information due their runaway status.[375]

Using multiple sources and methods, the identity and age verification can be strengthened.[376] After the initial age check, some platforms use additional tools for **age assurance**. The platform estimates if the age is correct based on what the children watch, how they talk and who they connect with. It should be noted that there is always a risk of overestimating the child's age and classifying them as older than they really are, as well as the risk of mistaking adults for children.[377] Even advanced and elaborate methods of age verification are being circumvented by children.[378]

Ways to strengthen this age verification after registration are through the use of **automatically generated data and user-reported data**. To start with the former, automatically generated data is information that is derived from a person's habits. This could be how they speak, locations they visit, or data that entered using a certain platform for example. Gathering this type of information puts together a profile of a person that can be used

[374] Nash, V., O'Connell, R., Zevenbergen, B. and Mishkin, A. (December 2012-December 2013) *Effective Age Verification Techniques: Lessons to Be Learnt from the Online Gambling Industry*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2658038 (Accessed 25 April 2022).

[375] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[376] Nash, V., O'Connell, R., Zevenbergen, B. and Mishkin, A. (December 2012-December 2013) *Effective Age Verification Techniques: Lessons to Be Learnt from the Online Gambling Industry*.

[377] Information Commissioner's Office (2021) *Age Assurance for the Children's Code*. Available at: https://ico.org.uk/media/about-the-ico/documents/4018659/age-assurance-opinion-202110.pdf (Accessed 25 April 2022).

[378] van der Hof, S. and Ouburg, S. (2021) *Methods for Obtaining Parental Consent and Maintaining Children Rights*. Available at: https://euconsent.eu/download/methods-for-obtaining-parental-consent-and-maintaining-children-rights/ (Accessed 25 April 2022).

to verify age.[379] Children will go to different locations and use different language for instance, which will give away their age. For user-reported data, different users of a certain platform can be asked to verify a person. This could, for instance, be an adult that the child selects, or other child users. A problem with this method is, however, that the bond between the child and the person that vouches for them is not necessarily verifiable, making it easier for children to circumvent this measure.[380]

It is crucial to adapt age verification tools to context specific challenges,[381] which is why deciding on the best strategies to protect children from OSEC through age verification requires careful analysis, including analysis of privacy implications. Other potentially effective methods include multi-factor identification, combining human inputs and technical verification,[382] and third-party age verification (e.g. ensured through background checks or tokenised age checks).[383]

At the same time, these measures need to be carefully examined for privacy concerns and inevitably generate more costs.[384] It is important to ensure that the use methods are privacy preserving, in line with the privacy by design and default principles, furthermore, sensitive personal data and automated profiling are to be avoided unless it is proved that it is in the best interests of the child.[385] General child rights impact assessment should be conducted to ensure that verification is proportionate, inclusive and children should participate in designing and developing methods of verification[386].

5. IMPLEMENT TECHNOLOGIES THAT IDENTIFY RISKS COMBINED WITH RISK MITIGATION STRATEGIES

RECOMMENDATION 5: Platforms should use computational risk detection to proactively identify potential risk and combine this with risk mitigation strategies, such as educational and awareness prompts and nudging, to help children be safer online.

SOLUTION IN DETAIL

- **Evidence-based computational risk detection combined with risk mitigation.** Computational risk detection can be an effective tool for making the online space safer for children, in particular in the case of grooming, sexual extortion, and unsafe sexting. To be effective, computational risk detection tools must be grounded in evidence and use a child-centred design approach. They should aim at identifying risks prior to, or in the moment they occur, and must be combined with age-appropriate risk mitigation strategies that use warning prompts and educational nudging to teach parents and children how to respond to risks.
- **Risk mitigation strategies including education/awareness design (e.g. warnings, prompts, nudges, and intelligent coaches)** (from service providers). Technologies can be used to help children make better decisions online, such as intelligent assistance or nudges supporting self-monitoring. As children are not likely to read or watch copious amounts of educational/training materials to learn how to be safe online, lightweight educational prompts offer the benefits of an in-the-moment educational approach in the context of a risky experience. A child, for instance, could get a warning and guidance on how best to proceed when contacted by a stranger. Such educational prompts must not be overly abstinence-based in terms of suggesting that young people not take any risks online, as children (especially teens) would then tune them out. Instead, they should guide children to appropriate action and provide tips as to how to cope with risks. Safety messaging should ideally be standardised across platforms rather than left up to the individual platforms. Risk mitigation strategies can also be effective to stop offenders and teenagers' risk

[379] Verification of Children Online (VoCO) project (2020) *Phase 2 Report*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/934131/November_VoCO_report_V4__pdf.pdf (Accessed 25 April 2022).

[380] Nash, V., O'Connell, R., Zevenbergen, B. and Mishkin, A. (December 2012-December 2013) *Effective Age Verification Techniques: Lessons to Be Learnt from the Online Gambling Industry*.

[381] Verification of Children Online (VoCO) project (2020) Phase 2 Report.

[382] Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls*.

[383] ibid

[384] ibid

[385] van der Hof, S. and Ouburg, S. (2021) *Methods for Obtaining Parental Consent and Maintaining Children Rights*.

[386] ibid

behaviour, for example, with the use of warnings if inappropriate content is about to be posted. Human-centred algorithms could be used to accurately detect online risks and trigger warnings or nudges.

RATIONALE

Technologies can be used to keep children safe by identifying risky situations where the children could be exposed to or involved in OSEC. Computational risk detection refers to machine learning and other automated technological tools that are deployed to detect risks. This can take the form of skin/nudity detection in pictures and videos, or use natural language processing.

Automated **skin detection** tools can be used to detect nudity and potentially unsafe and abusive material. Skin detection can prevent unsafe sexting and minimise exposure to CSAM and risks of sexual extortion, in particular if such tools are used together with risk mitigation approaches. Skin detection tools work by scanning images and videos using filtering systems such as light filters. Such filters identify patterns reflecting various skin tones for wavelengths of visible light.[388] Various models can be used where, for example, pixels most likely linked to non-skin areas of a material are discarded from a detected skin region of pixels and the region with detected skin is extracted for further processing.[389]

Many challenges still exist regarding this method of computational risk detection. Many tools are not fit to fully differentiate child and adult nudity. Risk detection models have struggled to identify children's faces because their facial traits are less defined.[390] Most of the tools are developed at software-level only, meaning that they can only apply to images and videos that have already been created and are most likely already disseminated. There is a lack of focus on designing such tools at hardware level (i.e. device level) to act as a preventative measure in case of sexting or OSEC. In addition, tools detecting nudity that are privacy-protecting use methods such as defocusing the image, which makes the detection less precise.[391]

Natural language processing is an AI tool that allows words and sentences to be identified and analysed using context clues, similar to how humans process language. There are already a wide variety of uses for natural language processing such as spell check, autocomplete, spam filters, and online assistants. Natural language processing tools can be deployed to identify potential grooming or sexual extortion. An algorithm can be developed and trained through a dataset of words and sentences linked to grooming using a machine learning method. The tool can calculate the probability that the messages translate into grooming attempts.[392] To be effective, such tools should also be accompanied by risk mitigation strategies such as educational prompts and nudging to help teenagers navigate the risk and seek support.

Platforms currently use predominantly scanning technologies such as PhotoDNA, which uses fingerprints of known CSAM to find duplicates on their platforms.[393] Their approach only identifies OSEC after it has occurred. More preventative approaches are needed to detect risks prior to, or as they occur, as well as combinations of such approaches with risk mitigation tools that help parents and teenagers learn to tackle the risk and ensure child online safety.[394]

[388] Tariq, M. U., Ghosh, A. K., Badillo-Urquiola, K., Jha, A., Koppal, S. and Wisniewski, P. (2018) 'Designing light filters to detect skin using a low-powered sensor', ***SoutheastCon 2018***, St. Petersburg, United States, 19-22 April. IEEE. **https://doi.org/10.1109/SECON.2018.8479027**.

[389] Tariq, M. U., Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2019) 'A Review of the Gaps and Opportunities of Nudity and Skin Detection Algorithmic Research for the Purpose of Combatting Adolescent Sexting Behaviors', ***21st International Conference on Human-Computer Interaction***. Orlando, United States, 26-31 July. Springer, Cham. **https://doi.org/10.1007/978-3-030-22636-7_6**.

[390] Bursztein, E., Clarke,E., DeLaune, M.,Elifff, D. M., Hsu, N., Olson, N., Shehan, J., Thakur, M., Thomas, K. and Bright, T. (2019) 'Rethinking the Detection of Child Sexual Abuse Imagery on the Internet', ***World Wide Web Conference*** (WWW '19). San Francisco, United States, 13-17 May. ACM, New York, United States, pp.2601–2607. **https://doi.org/10.1145/3308558.3313482**.

[391] Tariq, M. U., Ghosh, A. K., Badillo-Urquiola, K., Jha, A., Koppal, S. and Wisniewski, P. (2018) 'Designing light filters to detect skin using a low-powered sensor'.

[392] Muñoz, F., Isaza, G., Castillo, L. (2021) 'SMARTSEC4COP: Smart Cyber-Grooming Detection Using Natural Language Processing and Convolutional Neural Networks'. **https://doi.org/10.1007/978-3-030-53036-5_2**.

[393] Bursztein, E., Clarke,E., DeLaune, M.,Elifff, D. M., Hsu, N., Olson, N., Shehan, J., Thakur, M., Thomas, K. and Bright, T. (2019) 'Rethinking the Detection of Child Sexual Abuse Imagery on the Internet'.

[394] Tariq, M. U., Ghosh, A. K., Badillo-Urquiola, K., Jha, A., Koppal, S. and Wisniewski, P. (2018) 'Designing light filters to detect skin using a low-powered sensor'.

Similarly, in a systematic literature review of computational approaches for online sexual risk detection, Razi et al. (2021) found that the literature mostly proposed algorithms *"for detecting sexual predators (75%) after the sexual violence occurred (93%) using public datasets (82%)"*, many of which were not representative of the end users.[395] The literature also focuses on computational methods of detection with little on **risk mitigation strategies** once a risk is detected.[396] These findings indicate a need to develop tools that detect risks before victimisation occurs and that include risk mitigation strategies once detection occurs. **Teaching in the moment** constitutes an effective educational approach where the child may learn best how to tackle risk when faced with the risky situation, supported by risk detection and risk mitigation tools, such as generic warnings or prompts to reflect on their own behaviour.[397]

We know that teenagers benefit from learning how to cope and resolve risks in the moment through developing their own agency. Teaching children how to protect themselves is a key component of their developmental process.[398] Children as young as 7 years of age are able to identify online risks. Yet they need support as to how to deal with them.[399] Technological interventions could be designed to assist children in their experience of the online environment and, in particular, when they face risky situations or engage in risky behaviours. This could be done either through **educational prompts** when the child reaches out for support or wishes to have automated assistance. This could also be combined with risk detection interventions. The prompts could take children through a self-assessment of the risks and actions they can take to avoid the risks or cope with the risks. Alternatively, computational risk detection could, once a risk is detected, issue a **warning about the risk**, with an offer for additional information and support on how to deal with it.

Prompts can be effective in nudging children away from potential harmful behaviour. Educational prompts built into the design of apps and services can guide children and reinforce good behaviour.[400] Research demonstrates the benefits of warnings as a means of preventing abuse.[401] This approach tends to work better with older or more mature children. The impact of the educational prompts will vary depending on the maturity and competencies of children. The prompts must be adapted to the environment, type of use, and age of the child in order to engage them into changing their behaviour. Standard or inadequate messages might be ignored, especially if it seems like a 'parent telling you what to do'.[402]

Such interventions must be evidence-based and vetted by research in order to minimise incorrect flagging of risk and inadequate prompting. Prompts should be provided in a safe and privacy-protecting manner.

A review of the UK's Safety Technology sector highlights that businesses developed a range of technical solutions tackling online harms, such as threat detection and reporting, platform monitoring, takedown and domain alerts, URL lists, keyword collation and monitoring, hashing, content filtering, automated and human moderation, image processing, computer vision, and machine learning.[403] An example of technology

[395] Razi, A., Kim, S., Alsoubai, A., Stringhini, H., Solorio, T., De Choudhury, M. and Wisniewski, P. (2021) 'A Human-Centered Systematic Literature Review of the Computational Approaches for Online Sexual Risk Detection', *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 465, pp.1-38. **https://doi.org/10.1145/3479609**.

[396] Tariq, M. U., Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2019) 'A Review of the Gaps and Opportunities of Nudity and Skin Detection Algorithmic Research for the Purpose of Combatting Adolescent Sexting Behaviors'.

[397] Badillo-Urquiola, K., Chouhan, C., Chancellor, S, De Choudhary, M. and Wisniewski, P. (2019) 'Beyond Parental Control: Designing Adolescent Online Safety Apps Using Value Sensitive Design'.

[398] Pinter, A. T., Wisniewski, P., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Adolescent Online Safety: Moving Beyond Formative Evaluations to Designing Solutions for the Future'; Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewski, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online'; Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors'.

[399] Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewski, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online'.

[400] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[401] Prichard, J., Wortley, R., Watters, P.A., Spiranovic, C., Hunn, C. and Krone T. (2022) 'Effects of Automated Messages on Internet Users Attempting to Access "Barely Legal" Pornography', *Sexual Abuse,* 34(1), pp.106-124. **https://doi.org/10.1177/10790632211013809**.

[402] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[403] Donaldson, S., Davidson, J. and Aiken, M. (2021) *Safer technology, safer users: The UK as a world-leader in Safety Tech*. Available at: **https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/974414/Safer_technology__safer_users-_The_UK_as_a_world-leader_in_Safety_Tech_V2.pdf** (Accessed 27 April 2022).

developed by SafetoNet is presented below. The example is particularly relevant as it seems to combine risk detection and risk mitigation.

Figure 4. SafetoNet's automated risk detection approach

> SafeToNet offers an app that helps educate and safeguard children 'in-the-moment' when they use their device. It uses an intelligent keyboard that filters the child's outgoing messages to detect risk of harm in real time and provides feedback and advice to children as they type. This approach helps to prevent sexting, grooming, bullying, and abuse. The software also detects mental health issues such as dark thoughts, anxiety, and low self-esteem. SafetoNet guarantees that the child's rights to privacy are respected as no-one can see the messages sent or received.[404]

Another example of such an approach is the Online Well-Being and Safety (OWAS) Platform by Privately. OWAS software can be plugged into apps and games, and uses machine learning technology using natural language processing techniques, image analysis, and behavioural analysis techniques to identify risks and assist children in real time to help them deal with the risk. The technology respects privacy as personal data never leaves their device.[405]

One of the reasons online platforms have adopted a reactive approach to OSEC and have not embraced computational risk detection lies in the risk of being accused of interfering with the freedom of speech of their users and of over-policing of their online activities. Therefore, computation risk detection approaches need to be vetted and rooted in evidence-based research on how best to handle these situations (i.e. safety by design) in a way that is developmentally appropriate for children, depending on age and gender. Legal protections must also be adopted to ensure that children are not criminalised as a result of the use of risk detection tools. The tools and datasets should be transparent and available for research vetting.

Lastly, while algorithmic approaches are now standard technology, such interventions require a data protection and child rights impact assessment to ensure they meet the best standards in safety and respect children's rights.[406]

3.3 SAFETY DESIGNS IMAGINED BY CHILDREN

Children participating in the focus group discussions came up with safety designs and features themselves. The focus groups differed from each other in terms of the depth and creativity of what children discussed about safety designs. In some groups, the facilitators expressed that the children had difficulties thinking of features and needed some help.

In general, the children underlined the importance of already existing safety features, such as for privacy-related privacy and identity verification. They also gave some recommendations on what specific features for children could be and how the design could be more visible or more easily accessible. Overall, the children came up with **preventative, monitoring, and reactive measures** to ensure their safety online. Many of the children's ideas are in line with the above recommendations from the literature, including age verification, intelligent privacy features, risk detection, and parental control features adapted to children across all age groups.

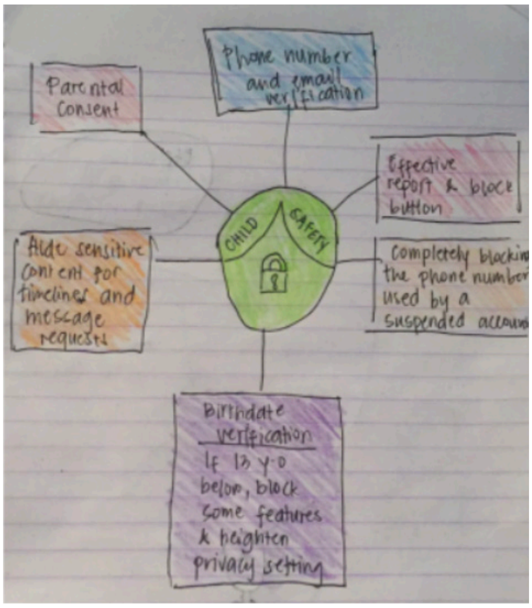
[404] idem – SafetoNet website available at: https://safetonet.com (Accessed 8 May 2022)
[405] Nguyen, S. (2018) *Applied case studies for children's data governance*. Available at: https://engagestandards.ieee.org/rs/211-FYL-955/images/IEEESA-Childrens-Data-Governance-Report.pdf (Accessed 8 May 2022).
[406] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

The following additional recommendations emerged from the focus group discussions, which were reflected in the literature to a less extent:

1. Introducing stronger reporting mechanisms and consequence for abusers
2. Increasing the visibility of rules and reporting mechanisms
3. Taking extra measures for content involving children
4. Using popular media and well-known people to deliver safety messages

Since many of the features that the children came up with exist to a certain extent, children are not proposing unrealistic features, but **call for action to** strengthen existing features and make some tweaks or additions. Some of the features that the children proposed are **relatively strict** and allow for **platforms and parents to have an active role** in keeping their child safe. This might counter platform creators' worries that children could feel monitored. The focus groups proved that children can see the added value of adults keeping them safe and have no problem with content being filtered or features being compromised for younger users. The biggest take-away might be that all the safety features should be more **visible**, so that it is **easy for children to take action** when they do feel unsafe. It is then perhaps not a case of finding new ways of keeping children safe, but rather of designing features it in such a way that they more appealing and attention-grabbing.

Figure 5. Safety features proposed by children


Parental Consent
Phone number and email verification
Effective report & block button
CHILD SAFETY
Hide sensitive content for timelines and message requests
completely blocking the phone number used by a suspended account
Birthdate verification
If 13 y-o below, block some features & heighten privacy setting

Source: Focus group in the Philippines

3.3.1 Children's design ideas in line with the literature

Children's ideas are presented below, starting with those in line with the literature, followed by additional proposals that are not reflected in the literature.

STRENGTHEN AGE AND IDENTITY VERIFICATION THROUGH OFFICIAL DOCUMENTS

To start with prevention, the children gave a variety of options on how to protect children before they are exposed to potentially harmful content or users. First of all, children need to be a certain age to be able to use social media platforms. In Nepal, this age was set to 13 years of age, as is also now the case for certain social media apps. In the Philippines, one participant proposed allowing children younger than 13 to use platforms provided there were additional safety features. In Romania, one of the children said:

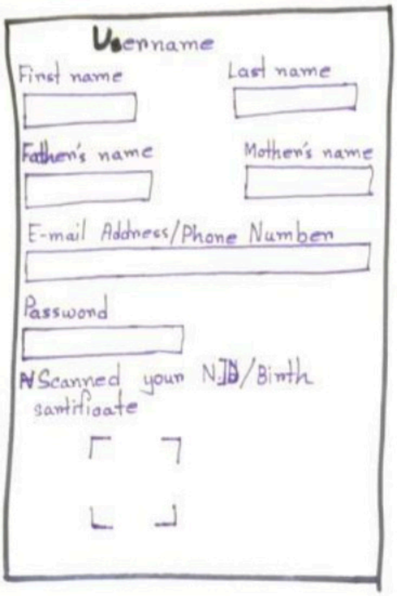
> *"I think it [an age requirement of 13] is okay, but I am also saying this because I am older than 13 years old. Otherwise, I would not agree that much with this."* (Focus group discussion, 14-year-old girl, Romania)

The girl was honest, in that she said she agreed with the age requirement because she herself had already surpassed the age. In various focus groups, some children also believed that younger children should be allowed on social media platforms. In the same focus group in Romania, the following was expressed:

> *"Well, it is my right to have access to social media."* (Focus group discussion, 13-year-old boy, Romania)

This was, however, often a minority or singular participant of the group. In general, the children indicated that access to social media platforms should not be guaranteed. The **age and identity** of all potential users should be checked before users are allowed on platforms. The children advised that a users' identity card or passport should be uploaded and checked. In the Netherlands, children expressed that anonymous accounts should no longer be allowed and that identity should be checked by providing a scan of the ID card. In one of the focus groups in Nepal, the children stated that parents should also check the identity of their children before the latter are allowed to use a platform. The parents' information could then be added to the child's profile, allowing notifications to be sent to the parents via email. Child participants in Bangladesh suggested asking for additional information from children, such as the name of their father and mother, and uploading a birth certificate as an alternative to an identity document.

Figure 6. Use of birth certificate for identity verification, proposed by children



Username

First name

Last name

Father's name

Mother's name

E-mail Address/Phone Number

Password

Scanned your NID/Birth sartificate

Source: Focus group in Bangladesh

DIFFERENTIATE FOR DIFFERENT AGE GROUPS AND USE INTELLIGENT SAFETY FEATURES

The children proposed special features designed for young people. The topic of **categorisation** (child-adult or even child-teen-adult) of users was proposed in multiple countries. In the Philippines, the children proposed an option where you could see the type of user an account belongs to, i.e. child, teen, or adult. This feature would allow children to use social media more safely, letting the child can control the 'feed' and content they can see. It is guided by bots that filter out friend requests from total strangers.

In Estonia, the participants suggested that there should be default safety features for children, allowing them to only see appropriate content and only connect with safe friends. In one of the focus groups in Nepal, a participant suggested that if it is established that a user is a child, their account should only appear to friends of friends and family. In the Philippines and Romania, it was even proposed to have separate platforms for children, or a **child-friendly version** like YouTube Kids, but in this case for example, Facebook Kids.

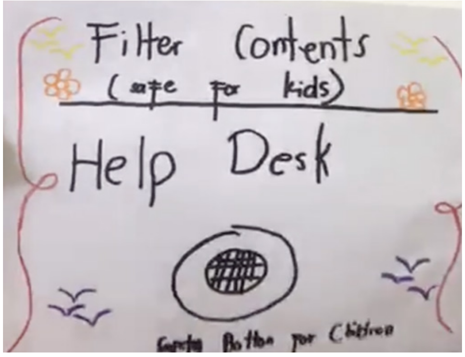
> *"If you have a younger age instead you should have more restrictions; there should be some platforms that are accessible only to children, not to adults."* (Focus group discussion, 15-year-old girl, Romania).

Other focus groups proposed that the platforms should filter the content children are able to see in their feed and through message requests. They suggested that settings be available to filter age-inappropriate content.

"I used to have an account [on Facebook] when I was younger and I got tagged in pornography pages, that is not okay." (Focus group discussion, 15 year old girl, Bolivia)

"Some contents on the Internet can affect the mental health of children, especially the contents about violence, sexual, and other sensitive contents" (Focus group discussion, 12-year-old girl, the Philippines).

Figure 7. Proposal by children to filter age-inappropriate content



Source: Focus group in the Philippines

INTERVENE IN POTENTIALLY HARMFUL SITUATIONS

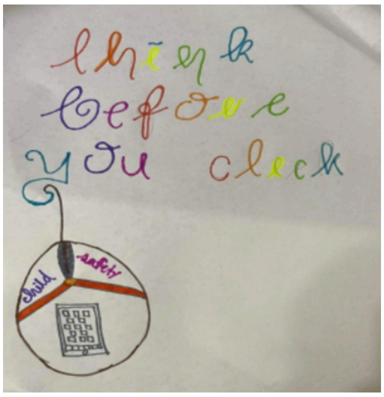
For monitoring, the children came up with measures that the platforms and parents could take. They indicated that the platform could use **moderators** who would oversee what content is shared and available to children and who could possibly intervene when the child was engaging in risky behaviour. This could range from deleting negative comments to intervening in the case of inappropriate interactions. Two children stated:

"TikTok should improve its design by not allowing bad comments, bad words, or insulting words to be posted on its application. For example, when someone wants to make a bad comment by calling another person 'dog' then, that word cannot be posted on the application." (Focus group discussions, Thailand)

"While watching videos on YouTube, negative videos popping up should be stopped."
(Focus group discussions, girl, Nepal)

In Romania, the children expressed that it could be beneficial if **pop-ups** appeared when accessing specific content or when sharing content. A similar idea came from one of the children in the Philippines, who suggested a 'think before you click' advertisement.

Figure 8. Example of pop-up safety reminder, proposed by children



Source: Focus group in the Philippines

In Bolivia, the children proposed a monitoring mechanism, where users that are not respecting the rules and regulations are followed, warned that they should not continue with similar messages, or can blocked.

PARENTS SHOULD BE INVOLVED IN KEEPING THEIR CHILDREN SAFE ONLINE

The children also said that parents should be able to have an overview of younger children's online behaviour (in Romania, they indicated 10 or 11 years of age), should have access to the account, or could choose what content their child is seeing. When children become older, parents should teach them more about online dangers. In one of the focus groups in Nepal, the children warned that parents should let their children use social media, as otherwise children will find a way to use it without informing their parents, which could be more dangerous.

In the grooming exercise that was discussed in section 2.1 under the sub-heading 'Interactions between children', children were also asked who should be the ones keeping them safe online. Parents were mentioned by almost every participant across all countries, making them the most mentioned actors. Other frequently mentioned actors were teachers, police officers, and social media platforms themselves.

> *"Parents play a big role for their children's safety".* (Focus group discussion, 15-year- old child, Bangladesh)

3.3.2 Additional design ideas proposed by children

The below design ideas imagined by children which were less reflected in the literature and are additional to the 6 solutions identified from the literature review.

STRONGER REPORTING MECHANISMS AND CONSEQUENCES FOR ABUSERS

SOLUTION 6: Reporting mechanisms should be strengthened with more options to support children being safe online. The children also emphasised that reporting should be enforced with faster responses from platforms and consequences for abusers.

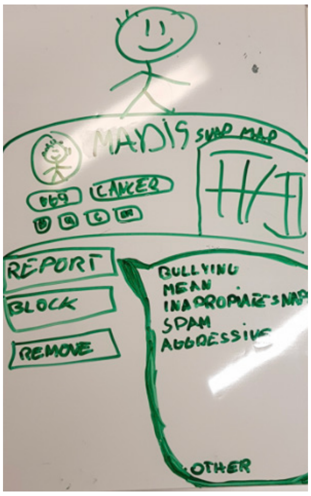
As reactive measures, the children stressed that existing measures such as **reporting, blocking users, and removing content** are good ways of ensuring children's safety online. In terms of reporting, the children asked for more awareness around reporting and more support when making a report. For instance, a **copy of the evidence** should be emailed directly to the police and to the owner of the platform. A platform should be designed in such a way to lower the threshold for reporting.

> *"If you have someone who is bothering you or approaches you with bad intentions, we should report it. There should be an easier way if I must do it."* (Focus group discussion, 15-year-old boy, Bolivia)

For the reporting mechanisms, children should be provided with various options to choose from. An example from the focus group in Estonia is provided below. In Bolivia, the children also emphasised that reporting should be enforced:

> *"Reporting should be faster. You report and nobody does anything."*
> (Focus group discussion, 13-year-old girl, Bolivia).

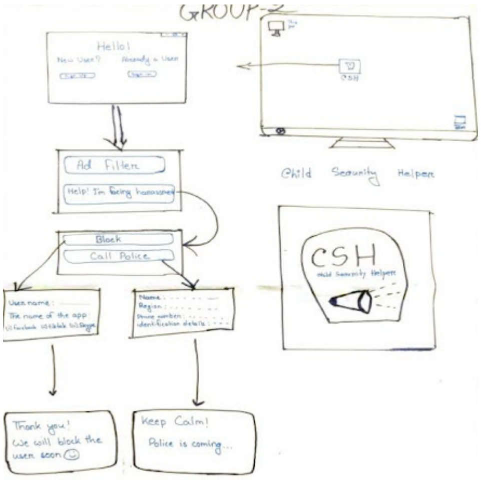
Figure 9. Children’s proposals for in-app reporting mechanisms



Source: Focus group in Estonia

The children in Bangladesh came up with a different reporting mechanism, called the ‘Child Security Helper’. This software would be launched after the young person signs up and has two functions: blocking all unwanted advertising and content, and providing help when someone is being harassed online. The latter would be in the form of a button present on the platform. When clicked, two options are provided: blocking the offending user and calling the police. To have someone blocked, the young person has to give their user name and the social media account they want to block. To call the police, the person has to give their name, address, phone number, and identity card.

Figure 10. Scheme of the functions of the ‘Child Security Helper’



Source: Focus group in Bangladesh

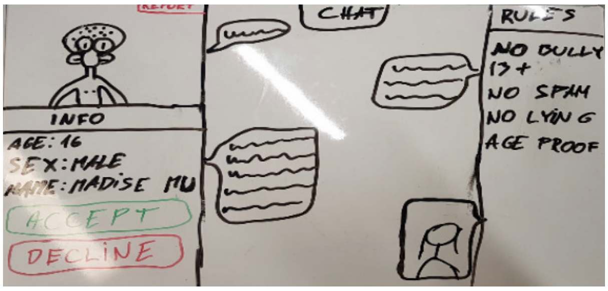
Children also proposed that platforms could take a faster approach by automatically removing certain content or words. In three focus groups, **consequences following violations** of the rules and regulations were also discussed. In the Philippines, it was proposed that every account should have a verified phone number. After a report that was deemed valid, the offending user would be banned and the associated phone number prevented from being used to register another account in the future. In Romania, the children proposed a system with minus points. If someone is reported, they receive a minus point and after a certain number of minus points, their profile is automatically deleted. Another idea from Romania was to have public warnings on the profiles of people that were abusive or aggressive to others.

CREATE MORE VISIBLE RULES, REGULATIONS, AND REPORTING MEASURES

SOLUTION 7: Children emphasised that rules and regulations should be more visible and that reporting measures need also more visibility. They suggested options to easily access emergency support or flag an online risk easily.

In five focus groups, children discussed **making rules and regulations more visible**. Some stated that more visibility would lead to more awareness among users about what is allowed and what is not. In Bolivia, the children proposed using advertising space for prevention messages. In focus groups in Nepal and Colombia, participants also stated that the terms and regulations of platforms should be more visible before children become members, and that children need to read these thoroughly. Others recommended more visibility for reporting measures, so that the threshold for reporting would become lower. In Bolivia and Colombia, children proposed having a 'danger button', or a visible aid that would facilitate the complaint and reporting procedure. One child in Estonia proposed a sidebar with the rules displayed, as shown in Figure 11:

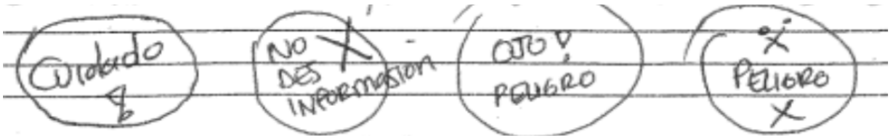
Figure 11. Proposal for a sidebar showing rules and regulations



Source: Focus group in Estonia

In Colombia, multiple children drew buttons with warnings, with one of the children adding different levels of danger, corresponding to 'Watch out', 'No disinformation', 'Warning, dangerous!', and ´Dangerous'.

Figure 12. Buttons showing different levels of danger



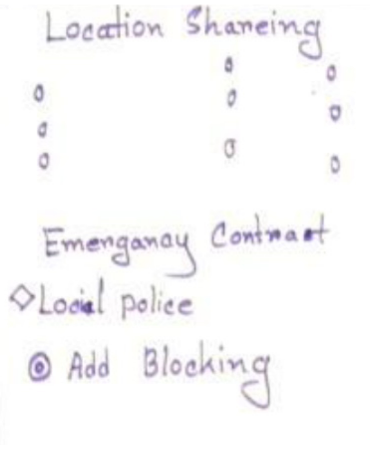
Source: Focus group in Colombia

Children in Romania suggested that social media platforms should enable users **to give the exact reason for blocking** someone by having an option to type in the reason(s) in detail and not only choose from a list of options. They also mentioned that the procedures for reporting someone take a lot of time and that social media platforms should have more staff responsible to deal with reported cases.

> *"It would be good if you could write the reason for giving a report without choosing a too restrictive category, to solve the problem immediately."* (Focus group discussions, 14-year-old girl, Romania).

In Bangladesh, the children also came up with a safety feature where children could press a button in the app which would allow them to **share their location** and **access emergency contacts** immediately. This would also come in handy when children are in physical danger at any time.

Figure 13. Button that shares the child's location in an emergency



Source: Focus group in Bangladesh

PLATFORMS TO ENSURE EXTRA PROTECTION FOR CONTENT INVOLVING CHILDREN

SOLUTION 8: Children advises that extra precautions should be taken to prevent children's information and content from being misused and from being disseminated without consent, such as preventing children's content to be downloaded or screenshotted.

Children thought that extra precautions should be taken to **prevent children's information and content from being misused**. Children shared concerns and even experiences of content they posted online being taken without their consent or knowledge.

> *"When a friend(s) called me via video call, then took a screenshot of my photos without my knowledge, then posted my photos on our (private) chat group."* (Focus group discussions, child, Thailand)
>
> *"[...] a person with no morals made screenshots of the pictures she uploaded and posted them on an illegal website saying she offered services to adults"* (Focus group discussions, 15-year-old girl, Bolivia).
>
> *"Because photos and videos of ourselves are personal, so if I would send them to someone, that person must really be my close friend. [...] Some may also do photoshopping/editing our photos and videos and distribute them widely to damage our reputations. [...] Some may also sell our photos, videos for earnings".*
> (Focus group discussions, child, Thailand)

A suggestion was to disable downloads or screenshots of information about children. In Bolivia, the children mentioned that downloading personal photos or screenshots should not be allowed.

> *"I think pictures one posts should be private and nobody should be able to download them or take screenshots of them."* (Focus group discussions, 15 years old girl, Bolivia)
>
> *"In google there is an incognito mode, and you can't take screenshots that way, it should be the same with social media apps."* (Focus group discussions, 15-year-old girl, Bolivia).

In Nepal, the children proposed that it should not be possible for content posted by children to go viral if it is about irrelevant and unnecessary topics. In several focus groups, children also recommended that people outside of their friend list should not be able to see children's profiles or send them messages, videos, photos, files, or links.

USE POPULAR MEDIA AND WELL-KNOWN PEOPLE TO DELIVER SAFETY MESSAGES

SOLUTION 9: Children suggested platforms to raise awareness about child safety online through safety messaging and online campaigns.

Another preventive recommendation provided by children is to raise **awareness about child safety online**. In Estonia and Colombia, the children proposed using **influencers** or *"people in showbusiness, people almost everyone knows"* (Focus group discussion, 13-year-old girl, Nicaragua) to deliver safety messaging, so that it will appeal to children. The message could be delivered through the platforms themselves, but the Colombian children also suggested using news programmes and television to disseminate the message.

> *"I would ask for help from an influencer so that they can help their followers."*
> (Focus group discussion, 15-year-old girl, Colombia)

Another girl in Nicaragua also proposed using **youth advocates**, such as herself, to spread awareness:

> *"I have more than 3000 followers because of the TikTok videos I make, so I can place the message there."*
> (Focus group discussion, 13-year-old girl, Nicaragua).

4. EU policy framework and recommendations

KEY FINDINGS

EU policy framework

- The European Union (EU) has **shared competence** with its Member States on matters relevant to OSEC
- The EU has **several legal bases** to adopt legislation tackling OSEC, including Articles 82 and 83 of the Treaty on the Functioning of the European Union (TFEU) and the internal market legal basis
- Some key legislation has already been adopted, such as the Child Sexual Abuse Directive, the e-Commerce Directive, and the Interim Regulation
- EU legislation already includes some safety by design requirements in the **Audio-visual Media Service Directive** (AVMSD), including transparent and user-friendly reporting mechanisms, age verification systems and parental control
- The AVMSD also requires video-sharing platforms to ensure that children may not 'normally hear or see' content that may impair their physical, mental or moral development. It does not, however, define what type of content fall in this category
- The requirements are **largely insufficient** to tackle OSEC, and the scope of the AVMSD is limited to video-sharing and video-on-demand platforms
- The General Data Protection Regulation (GDPR) sets consent for data processing at 13 years old and advocates a data protection by design approach

KEY RECOMMENDATIONS TO THE EU

- **Decriminalise** the exchange of intimate content and material among children in its revision of the Child Sexual Abuse Directive
  - Consensual exchange of intimate content and material among children to be decriminalised
  - Non-consensual exchange of intimate content and material among children must be clearly defined
- Require platforms to have **effective age verification** systems
  - Require the use of age verification methods for all platforms
  - Add requirements for effective, privacy-preserving age assurance
  - Explore possibilities to help online platforms verify age and identity
- Require all platforms to have **transparent reporting mechanisms** and **referral** systems for children reporting OSEC
  - Clarify and strengthen the transparency requirements of online reporting in the Directive
  - Require platforms to refer children to appropriate help services after they make an online report
  - Establish mandatory reporting mechanisms in the event that a child reports a form of OSEC
- Make online platforms legally accountable for **minimum safety standards** to keep children safe
  - The EU should make platforms accountable for establishing minimum safety standards
  - The minimum standards should include a requirement for a child rights impact assessment (CRIA)
- Support and strengthen initiatives aimed at **education, awareness, action research,** and **offender interventions**
  - Initiate and fund education and awareness programmes
  - Fund safe, online support platforms
  - Fund action research on OSEC and CSAM
  - Fund interventions aimed at OSEC offenders

4.1 EU LEGAL AND POLICY FRAMEWORK ON OSEC: PROTECTION AND GAPS

The European Union (EU), representing 27 Member States, is in a position to set global standards to protect children from online harm, including OSEC. The EU has the ambition to fight online child sexual abuse and has already established certain requirements and has some key proposed legislation in the pipeline. However, the EU can only act insofar as its competences allow, in line with the principle of subsidiarity, as laid down in the **Treaty on European Union (TEU)**. This section provides an overview of the EU competences and its legal framework, and highlights gaps that exist.

4.1.1 EU competences on OSEC and child safety by design

The EU has a range of tools and actions it can take to tackle online child sexual abuse and exploitation. However, it is limited in its scope of action by the Treaties that grant its power.

Protecting the rights of the child is one of the EU's main aims, as laid down under Article 3 of the TEU. This means that the EU, its institutions and its Member States have an obligation to promote, protect, and fulfil the rights of the child. Additionally, Article 24 of the **Charter of Fundamental Rights of the EU** establishes the obligation of the EU and Member States to protect the rights of the child when implementing EU law.[407] Accordingly, public authorities and private institutions must have the child's best interests as a primary consideration in all actions relating to children.

The EU has three different categories of competences where it is authorised to act via the EU Treaties in various areas: exclusive, shared and supporting competences. Additionally, in some areas the EU has a special competence, where it is allowed to play a particular role or to go beyond what it is normally allowed under the Treaties.[408] In sum, the EU may legislate only where it has been given competence under the TEU and the TFEU.

The EU has shared competence on matters falling under the scope of justice and fundamental rights, which includes the rights of the child. Shared competence means that both the EU and its Member States may adopt legally binding acts in this area. In the areas of education, vocational training, youth and sport, and civil protection, the EU has a supporting competence, which means the EU can only intervene to support, coordinate, or complement the action of EU countries. Importantly, in the case of a supporting competence, the EU may adopt legally binding acts but must not, however, require the harmonisation of Member State laws and regulations.

Regarding children's rights as a whole, there is no specific competence to legislate as a general area. However, as children's rights is a cross-sectorial field, EU competence varies based on the specific issue at hand. Areas relevant for children's rights where the EU has extensively legislated are: data and consumer protection, asylum and migration, and cooperation in civil and criminal matters.

The TEU and TFEU provide for a legal basis for EU action on child sexual exploitation and abuse. The Treaties have enhanced the EU's capacity to adopt binding legal instruments on child protection related to criminal offences and procedures, however these are limited to the aspects identified in Article 81 TFEU for matters affecting children in judicial cooperation and civil matters, Article 82 TFEU for matters affecting children in judicial cooperation in criminal matters and Article 83 TFEU for the establishment of minimum rules, criminal offences, and sanctions in various areas including the sexual exploitation of children. These provisions have led to the adoption of various directives related to trafficking and the sexual exploitation of children, such as Directive 2011/93/EU and its revision (see Section 3.2). Moreover, Article 216 TFEU enables the EU to conclude international conventions in relation to children's rights; the EU has not concluded the UNCRC as an entity, while all its Member States have done so.

[407] European Commission (n.d.) *EU action on the rights of the child*. Available at: **https://ec.europa.eu/info/policies/justice-and-fundamental-rights/rights-child/eu-action-rights-child_en** (Accessed 18 April 2022).

[408] Consolidated version of the Treaty on the Functioning of the European Union (TFEU) (2016) *Official Journal of the European Union* C 202/1, Articles 2-6. Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12016E/TXT** (Accessed 29 April 2022).

The EU has shared competence on matters falling under the scope of justice and fundamental rights, which includes the rights of the child.[409] Shared competence means that both the EU and its Member States may adopt legally binding acts in this area. In the areas of education, vocational training, youth and sport, and civil protection, the EU has a supporting competence, which means the EU can only intervene to support, coordinate, or complement the action of EU countries. Importantly, in the case of a supporting competence, the EU may adopt legally binding acts but must not, however, require the harmonisation of Member State laws and regulations.

Regarding children's rights as a whole, there is no specific competence to legislate as a general area. However, as children's rights is a cross-sectorial field, EU competence varies based on the specific issue at hand. **Areas relevant for children's rights where the EU has extensively legislated are: data and consumer protection, asylum and migration, and cooperation in civil and criminal matters**.[410]

The TEU and TFEU provide for a legal basis for EU action on child sexual exploitation and abuse. The Treaties have enhanced the EU's capacity to adopt binding legal instruments on child protection related to criminal offences and procedures, however these are limited to the aspects identified in Article 81 TFEU for matters affecting children in judicial cooperation and civil matters, Article 82 TFEU for matters affecting children in judicial cooperation in criminal matters and Article 83 TFEU for the establishment of minimum rules, criminal offences, and sanctions in various areas including the sexual exploitation of children. These provisions have led to the adoption of various directives related to trafficking and the sexual exploitation of children, such as Directive 2011/93/EU and its revision (see Section 3.2). Moreover, Article 216 TFEU enables the EU to conclude international conventions in relation to children's rights;[411] the EU has not concluded the UNCRC as an entity, while all its Member States have done so.

Criminal law legal bases (Articles 82 and 83 TFEU) provide avenues for the EU to legislate on OSEC. The EU has the power to *"progressively adopt measures establishing minimum rules relating to the constituent elements of criminal acts and penalties in the fields of organised crime, terrorism, and illicit drug trafficking"*.[412] This competence can be used to legislate on the topic of online sexual exploitation of children. Directive 2011/93/EU was adopted based on Articles 82(2) and 83(1) of the TFEU along with Article 6(1) TEU.

The EU may legislate from a consumer angle regarding OSEC. For instance the e-Commerce Directive is based on the internal market legal basis (Article 114 TFEU) and includes a requirement that internet service providers remove illegal material once they are made aware of such material's presence on their platforms. As such, the current EU framework in governing the liability of intermediaries is cross-industry, as it includes the regulation of the technology sector, private sector, and law enforcement, and uphold fundamental rights. Additionally, governing the liability of online platforms as intermediaries must be in accordance with other EU policies and legal frameworks.

The technology sector is regulated under the EU's shared competence, including the regulation of key technologies, technological developments, and corporate social responsibility. Additionally, due to the established fact that sexual exploitation and abuse constitute serious criminal matters falling under the EU's competence on criminal matters, intermediary criminal liability may be regulated under the EU legal framework.

[409] European Commission (n.d.) ***Areas of EU action***. Available at: **https://ec.europa.eu/info/about-european-commission/what-european-commission-does/law/areas-eu-action_en#EU** (Accessed 18 April 2022).

[410] European Parliament (2012) ***EU Framework of Law for Children's Rights***. Available at: **https://www.europarl.europa.eu/RegData/etudes/note/join/2012/462445/IPOL-LIBE_NT(2012)462445_EN.pdf** (Accessed 18 April 2022).

[411] Odink, I. (2019) ***Children's Rights in the EU: Marking 30 Years of the UN Convention on the Rights of the Child. European Parliamentary Research Service***. Available at: **https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/644175/EPRS_BRI(2019)644175_EN.pdf** (Accessed 21 April 2022).

[412] Consolidated version of the Treaty on European Union (TEU) (2016) ***Official Journal of the European Union C 202/1***, Article 31(1)(e). Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A12016M%2FTXT** (Accessed 29 April 2022).

4.1.2 EU framework on child exploitation and abuse

Under the premise of promoting, protecting, and fulfilling the rights of children, the EU has made the fight against child sexual exploitation a top priority. The European Commission adopted the **EU strategy for 2020-2025 for a more effective fight against child sexual abuse**, which aims to provide a framework for developing a strong and comprehensive response to child sexual abuse and exploitation, in both online and offline forms.[413] The strategy has eight main initiatives. Priority five is to *'"enable Member States to better protect children through prevention"'*. The Strategy focuses on prevention programmes for offenders and awareness raising, notably of parents and children. The latter forms part of the safety by design approach, notably in empowering children to speak up, react, and report. Yet, the Strategy falls short of embracing safety by design. Priority seven aims to *"galvanise industry efforts to ensure the protection of children in their products"*. This priority focuses on detection and reporting of CSAM rather than designing platforms that are safe for children in the first place.

The main legal instrument concerning online child exploitation and abuse is the **Child Sexual Abuse Directive**.[414] The Directive requires EU Member States to take all necessary necessary to ensure that the production, acquisition, possession, distribution, dissemination, transmission, offering, supplying, or making available of CSAM, or merely knowingly obtaining access to CSAM is punishable under national laws. The Directive was the first comprehensive EU legal instrument establishing minimum rules concerning the area of child sexual abuse and exploitation of children, both in online and offline domains.

As announced in the Strategy, the Child Sexual Abuse Directive will be revised. In addition to the implementation of the Directive as a matter of priority,[415] the Commission wants a revision of the directive to tackle legislative gaps and adapt the text to the technological developments that have been made in the last decade. The latter has resulted in an increase of CSAM being circulated and the emergence of new forms of child sexual abuse and exploitation online due to both new social media and wider internet access. In 15 years (2005-2020), the amount of potential CSAM material reported to the National Center for Missing and Exploited Children (NCMEC) increased by 15,000%.[416] In 2020 alone, *"1.038.268 content URLs were entered into ICCAM"*, the INHOPE platform used by 47 hotlines (30 of which are based in the EU).[417] In total, 77% of the identified CSAM depicted children below the age of 13 years old.[418] Until the early 2000s, OSEC consisted mostly of material, such as CSAM, that was circulating. The emergency of new technologies and new ways to communication online has brought new forms of OSEC, such as online grooming, live streaming of child sexual abuse, sexual extortion of children, self-generated material involving children, and deepfakes, among others.[419]

From a consumer protection point of view, the **e-Commerce Directive**[420] establishes the liability of internet service providers for illegal activity and/or content on their services (Article 14). It requires internet service providers to expeditiously remove or to disable access to the illegal activities, including OSEC, upon obtaining knowledge or awareness. However, the e-Commerce Directive does not provide a general obligation to actively monitor or seek facts or circumstances indicating illegal activity and report those to law enforcement.

[413] European Commission (2020) *EU strategy for a more effective fight against child sexual abuse*. Available at: https://ec.europa.eu/home-affairs/system/files/2020-07/20200724_com-2020-607-commission-communication_en.pdf (Accessed 22 April 2022).

[414] Directive (EU) 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA (2011) *Official Journal of the European Union* L 335. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02011L0093-20111217 (Accessed 21 April 2022).

[415] European Commission (2020) *EU strategy for a more effective fight against child sexual abuse*.

[416] Thorn (2020) *The road to Safer: Equipping industry to end CSAM*. Available at: https://www.thorn.org/blog/announcing-safer-built-by-thorn-eliminate-csam/ (Accessed 21 April 2022).

[417] INHOPE (2021) *Annual Report 2020*. Available at: https://inhope.org/media/pages/the-facts/download-our-whitepapers/annual-report/bb4dd3cdc3-1628156678/inhope-annual-report-2020.pdf (Accessed 18 April 2022).

[418] ibid

[419] ECPAT International (2020) *Summary Paper on Online Sexual Exploitation of Children*. Available at: https://ecpat.org/wp-content/uploads/2021/05/ECPAT-Summary-paper-on-Online-Child-Sexual-Exploitation-2020.pdf (Accessed 21 April 2022).

[420] Directive (EU) 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce') (2000) *Official Journal of the European Union* L 178. Available at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32000L0031 (Accessed 21 April 2022).

In 2021, the European Commission adopted an interim Regulation to ensure that online communications service providers could continue their voluntary practices to detect, remove, and report CSAM. The **Interim Regulation** contains a temporary derogation from certain provisions of the ePrivacy Directive for the purpose of combatting child sexual abuse online,[421] notably from Articles 5(1) and 6 of the ePrivacy Directive protecting the confidentiality of communications and traffic data. The temporary derogation enables interpersonal communications services providers to continue using specific technologies to detect, report, and remove CSAM. The interim derogation was necessary due to the original wording of the ePrivacy Directive which could have prevented the use of hashing technologies to automatically detect previously identified CSAM hosted on service platforms, thereby posing a legal obstacle for CSAM detection in Europe. Voluntary efforts by platforms to detect online grooming may continue through the use of already existing technologies.

In addition, the 2011 **EU Anti-trafficking Directive** does not provide any guidance on the liability of online platforms. The EU has launched a public consultation on the implementation of the Anti-trafficking Directive to assess the need for its revision, and the extent to which it should be revised. One of the main concerns raised is the need to tackle new forms of trafficking, including the increased use of online platforms.

Currently, there are several legislative proposals concerning child exploitation and abuse: the proposal for a regulation on child sexual abuse online, the proposal for a directive on online gender-based violence, the proposed **Digital Service Act** as well as the revision of **Directive 2011/93/EU**. The proposal for a regulation on child sexual abuse online aims to set out the responsibilities of relevant online service providers, including the requirement to detect and report CSAM to public authorities.[422] In addition, the proposed regulation will also explore the creation of a European centre to prevent and counter (online) child sexual abuse. The proposal for a directive on online gender-based violence aims to provide uniform standards for preventing this kind of violence, including the combatting of illegal and harmful gendered online content.[423] The proposed Digital Service Act aims to create a safer and trusted online environment by establishing a framework of layered responsibilities targeted at different types of services such as online platforms.[424] The proposed Act aims to introduce EU-wide obligations to ensure transparency, accountability, and regulatory oversight in the digital world, to ultimately protect children and young people online, amongst others. Lastly, the Child Sexual Abuse Directive and its implementation is being assessed and to identify its legislative gaps, best practices, and priority actions at EU level.[425]

4.1.3 Safety by design in the EU legal framework

The term safety by design is not reflected in EU legal frameworks, however elements can be identified in some legislative texts. In this Section, we look at both physical products and online services in order to draw a comparison between the approaches taken.

EU LEGISLATION ON OFFLINE PRODUCTS AIMED AT CHILDREN

The EU adopted a legal framework related to the safety of physical toys marketed at children in the EU as far back as 1988 with the first Toy Safety Directive, which has since then been updated.

The 2009 **Toy Safety Directive**[426] sets minimum safety requirements toys need to meet in order to be allowed in the EU market. The Directive concerns toys designed or intended for play by children under the age of 14.

[421] European Commission (n.d.) *Legal framework to protect children*. Available at: https://ec.europa.eu/home-affairs/policies/internal-security/child-sexual-abuse/legal-framework-protect-children_en (Accessed 22 April 2022).

[422] European Commission (n.d.) *Fighting child sexual abuse: detection, removal and reporting of illegal content online*. Available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12726-Fighting-child-sexual-abuse-detection-removal-and-reporting-of-illegal-content-online_en (Accessed 21 April 2022).

[423] European Commission (n.d.) Combating gender-based violence – protecting victims and punishing offenders. Available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12682-Combating-gender-based-violence-protecting-victims-and-punishing-offenders_en (Accessed 21 April 2022).

[424] European Parliament (n.d.) *Legislative Train Schedule – A Europe Fit for the Digital Age*. Available at: https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-digital-services-act (Accessed 21 April 2022).

[425] European Commission (n.d.) *Combating child sexual abuse – review of EU rules*. Available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13073-Combating-child-sexual-abuse-review-of-EU-rules_en (Accessed 21 April 2022).

[426] Directive (EU) 2009/48/EC of the European Parliament and of the Council of 18 June 2009 on the safety of toys (2009) *Official Journal of the European Union* L 170. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02009L0048-20210521 (Accessed 22 April 2022).

The requirements address physical health and safety risks in general as well as a list of particular risks identified in the Directive (Articles 4 and 10, and Annex II). The Directive also requires manufacturers to carry out safety assessments before a toy reaches the EU market, including in relation to the risk of potential exposure to chemical, physical, mechanical, electrical, and hygiene hazards (Article 18).

The manufacturer must demonstrate the compliance of a toy by self-verification by using the European harmonised standards and, in certain cases, by third-party verification through a notified body.

In short, the Toy Safety Directive provides minimum standards to ensure the basic safety of the toys with the goal of protecting children from harm. While the understanding of harm under the Directive remains limited, it provides protection against the most serious health risks. Similar protection does not yet exist for online games and play. Lastly, the Directive embodies some aspects of child safety by design (notably accountability and transparency) by requiring that toys are designed in accordance with the health and safety of children, and that manufacturers carry out sample testing of marketed toys, investigate complaints, recall toys if necessary, and keep distributors informed of any issue. The scope of the requirements remain too limited to qualify as safety by design requirements.

The EU is set to amend the Toy Safety Directive in 2022. To that end, 5Rights has called on the European Commission to ensure "*that digital toys, including connected toys and standalone software (games, social media and other apps)*" are explicitly in scope.[427]

The **General Product Safety Directive**[428] provides a wide framework for safe products on the EU market. It applies when no specific legislation or provision exists to regulate specific products (such as the Toy Safety Directive) or where aspects and risks, or categories of risks, are not covered by specific requirements. It covers both products and services intended for consumers or likely to be used by consumers in the course of a commercial activity, and applies equally to products sold online. The General Product Safety Directive requires producers to place only safe products on the market (Article 3) and to provide relevant information to enable consumers to assess the risks inherent in a product and to take precautions against those risks (Article 5). Similarly to the Toy Safety Directive, the principles of safety by design are not fully embodied in the legislation.

The European Commission has adopted a proposal to replace the General Product Safety Directive by a Regulation[429] clarifying, among other things, the application of the Directive of online marketplaces. It also makes clear that the Regulation would not include a general obligation for online marketplaces to monitor their platform or to actively identify illegal activity. Online marketplaces would only be required to expeditiously remove content referring to dangerous products, upon obtaining actual knowledge or awareness of the illegal content.

EU LEGISLATION REGULATING ONLINE SOCIAL MEDIA

The EU legal framework provides some basic safety requirements applicable to online social media to certain extent.

[427] 5Rights Foundation (2021) T*oy Safety Directive Review*. Available at: https://5rightsfoundation.com/uploads/ToySafetyDirectiveReview-5RightsRecommendations.pdf (Accessed 25 April 2022).

[428] Directive (EU) 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety (2001) *Official Journal* L 011. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02001L0095-20100101 (Accessed 22 April 2022).

[429] European Commission (2021) *Proposal for a Regulation of The European Parliament and of The Council on general product safety, amending Regulation (EU)* No 1025/2012. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0346&qid=1628522210573 (Accessed 22 April 2022).

[430] Directive (EU) 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) (2010) *Official Journal L* 095. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02010L0013-20181218 (Accessed 22 April 2022).

The **Audiovisual Media Services Directive** (AVMSD)[430] provides some requirements related to the safety of children online. It applies to audio-visual media on video-sharing as well as on-demand platforms[431] (thereby include apps such as TikTok). The AVMSD aims to establish an EU single market of audiovisual services and thus to ensure common rules and key principles such as freedom of retransmission of services from other Member States, prohibition of hate speech and incitement to hatred, as well as protection of minors from harmful content.

Under the AVMSD, (broadcasting) media service providers must ensure that children may not 'normally hear or see' content that may **i**mpair their physical, mental or moral development. The Directive does not prescribe specific measures but suggests that this could be accomplished through selecting the time of the broadcast, age verification tools, or other technical measures. The measures must be proportionate to the potential harm of the programme, meaning that stricter measures should apply to the most harmful content such as violence and pornography (Article 6a). The AVMSD also requires that media service providers inform viewers about the potentially harmful nature of content that may impair the physical, mental or moral development of children. However, the AVMSD does not define what constitutes harmful content.

In terms of video-sharing platforms, similar provisions apply requiring that video-sharing platform providers take appropriate, practicable and proportionate measures to protect children from *"programmes, user-generated videos and audiovisual commercial communications which may impair their physical, mental or moral development"* (Article 28b).

The AVMSD provides a list of measures to protect children that platforms must comply with 'as appropriate'. The list includes safety by design features, such as:

- transparent and user-friendly reporting mechanisms for users to report or flag content to the video-sharing platform provider;
- systems through which users receive an explanation of the outcome of the reporting and flagging;
- establishing "*age verification systems for users of video-sharing platforms with respect to content which may impair the physical, mental or moral development of minors"*;
- easy-to-use systems allowing users to rate content that may impair the physical, mental or moral development of minors;
- parental control systems that are under the control of the end user with respect to content that may impair the physical, mental or moral development of minors;
- transparent, easy-to-use and effective procedures for handling users' complaints; and
- providing effective media literacy and awareness raising of those measures and tools (Article 28b).

The AVMSD provides some safety by design requirements. However, these are limited. In relation to age verification systems for instance, the only requirement is that such a system is in place. There is no requirement for it to be effective. We know that systems currently in use are easily circumvented by children (see Section 3). The EUConsent legal landscape study[432] reports that national laws have not established criteria or further guidance for the implementation of effective age verification systems in light of the absence of strong EU requirements. In addition, a definition of what constitutes content *"which may impair the physical, mental or moral development of minors"* is missing. Arguably, at the minimum, serious forms of OSEC would fall within such a category.

[431] Under the Directive, 'video-sharing platform service' is defined as a service *"where the principal purpose of the service or of a dissociable section thereof or an essential functionality of the service is devoted to providing programmes, user-generated videos, or both, to the general public, for which the video-sharing platform provider does not have editorial responsibility [...] and the organisation of which is determined by the video-sharing platform provider, including by automatic means or algorithms in particular by displaying, tagging and sequencing"*.

[432] Caglar, C. and Nair, A. (2021) *EU Member State Legal Framework*. Available at: https://euconsent.eu/download/eu-member-state-legal-framework/ (Accessed 22 April 2022).

Lastly, the AVMSD does not specifically regulate social media platforms. It only covers platforms for which user-generated videos constitute an *"essential functionality of the service"* (Article 1). According to Recital 5 of Directive 2018/1808/EU, audiovisual content must not be "merely ancillary to, or a minor part of" the service. It is left to national authorities to assess whether a service falls under that category depending on *"the nature and the particular role played by user generated videos"* in the service offered by the platform, taking into account qualitative and/or quantitative considerations.[433] It is likely that Instagram, TikTok and Snapchat would fall under this category.

EU DATA PROTECTION

Data protection and privacy concerns in the EU are governed by the **General Data Protection Regulation** (GDPR). The GDPR is applicable everywhere where data related to persons within the EU is targeted or collected. The GDPR pushes for data protection by design and default. The GDPR tackles the child's consent in Article 8, which defines a child as anyone below the age of 16, however, Member States may lower this age provided that this is not lower than 13 years of age. Data controllers who fall under the definition provided in the GDPR should make reasonable efforts to verify the consent given by the child or authorisation given by parents and/or legal guardians, taking into consideration the available technology. Recital 38 of the GDPR states that children require 'specific protection' with regard to their personal data, as they may be less aware of the risks, consequences and safeguards and their rights in relation to the processing of their personal data.[434]

4.1.4 Summary overview of the EU framework and proposals

The Table below provides an overview of the wide range of EU legislation that currently addresses child safety from various angles. The Audiovisual Media Services Directive is possibly the legislation that encompasses the most safety by design requirements; yet falls short from adequately addressing online risks faced by children. Many of the below pieces of legislation are set to be revised. It would thus be the opportunity, through those revisions, for the adoption of legal requirements that embrace effective child safety by design.

Table 5. Overview of the current EU legal framework and proposals

Child sexual exploitation / criminal law	Digital environment / platforms	Product safety	Privacy / data protection
Child Sexual Abuse Directive – to be revised	e-Commerce Directive	Toy Safety Directive – to be revised	GDPR
EU Anti-trafficking Directive – to be revised	Interim Regulation allowing voluntary detection	General Product Safety Directive – to be revised	ePrivacy Directive
Proposal for a Regulation on fighting child sexual abuse: detection, removal and reporting of illegal content online	Audiovisual Media Services Directive		
Proposal for a Directive on combatting violence against women and domestic violence (including online)	Proposal for a Digital Service Act		

[433] 'Communication from the Commission Guidelines on the practical application of the essential functionality criterion of the definition of a 'video-sharing platform service' under the Audiovisual Media Services Directive (2020/C 223/02)' (2020) *Official Journal of the European Union* C 223/3. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C_.2020.223.01.0003.01.ENG&toc=OJ:C:2020:223:TOC (Accessed 22 April 2022).

[434] Regulation (EU) 2016/679 of The European Parliament and of The Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (2016) *Official Journal of the European Union* L 119/1, Recital 38. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=DA (Accessed 22 April 2022).

4.2 EU POLICY RECOMMENDATIONS

The EU policy framework already establishes some minimum standards for fighting OSEC, yet EU legislation is largely insufficient to tackle the scale and complexity of the issue. It does not sufficiently incentivise social media platforms to put in place effective safety measures and the legal provisions are too vague to lead to effective safety measures.

The EU framework has incorporated some elements of safety by design in its requirements, including transparent and user-friendly reporting mechanisms, age verification and parental control systems. The Audio-visual Media Service Directive (AVMSD) a key instrument in this area. Yet, it is limited in scope and lacks effective requirements.

The EU has the power to enact stronger policies and, in representing 27 Member States, is well positioned to adopt global standards that can protect children worldwide.

The below policy recommendations have been selected based on the insights from this research, including the literature review, focus groups with children, and inputs from the senior experts and international experts who participated in the online workshop.

4.2.1. EU to decriminalise exchange of intimate content and material among children in its revision of the Child Sexual Abuse Directive

Recommendation 1: Children can be part of the OSEC problem through engaging in at risk behaviours or even perpetrating OSEC themselves. Yet, in most cases, a criminal response is not adequate to address the issue. The line between consensual sexting and risk behaviours or even OSEC is blurry. Currently, the legislation does not reflect the complexity of the phenomenon of sexting. The EU must also ensure that children can safely explore their sexuality without fear of a criminal response by decriminalising consensual exchange of intimate content between children and defining consensual and non-consensual exchange of intimate content.

CALL FOR ACTION

- **Consensual exchange of intimate content and material among children to be decriminalised.** The EU should clarify in its revision of the Child Sexual Abuse Directive that the consensual exchange of intimate content and material (sexting) among minors is not a criminal offence. It should adopt a clear definition of consensual (legal) sexting that is not linked to the age of consent, i.e. the age at which a person is considered to be legally competent to consent to sexual acts, and which differs by country.
- **Non-consensual exchange of intimate content and material among children must be clearly defined.** EU law should adopt a clear definition of non-consensual sexting that amounts to a criminal offence, which covers situations where a child is pressured or coerced into producing and sharing intimate content, as well as the non-consensual further distribution of intimate content that was originally shared with consent.

RATIONALE

Sending sexually explicit material/sexting is becoming the norm among minors and is part of their exploration of sexuality. Teenagers have sexual interactions online with peers. In a global survey, 33% of the respondents had received sexually explicit content from peers.[435] Adolescents use sexting to take a relationship further and this can be seen as a natural progression of a romantic relationship, when consensual.[436]

[435] Economist Impact and WeProtect Global Alliance (2022) ***Estimates of childhood exposure to online sexual harms and their risk factors***. Available at: https://www.weprotect.org/economist-impact-global-survey/#report (Accessed 18 April 2022).

[436] Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2020) 'Let's Talk about Sext: How Adolescents Seek Support and Advice about Their Online Sexual Experiences', ***2020 CHI Conference on Human Factors in Computing Systems***, Honolulu, United States, 25-30 April. ACM, New York, United States. https://doi.org/10.1145/3313831.3376400.

However, sexting is associated with risks of online violence, including further dissemination of material without consent. Children may feel pressured to participate in sexting, leading to negative experiences. Children report it difficult to decline requests for sexting and unsolicited sexts that come from peers, partners, or friends of mutual contacts.[437] Research also confirms that the negative impact of unsolicited sexts leads to higher depression and anxiety, and lower self-esteem.[438]

Currently, the legislation does not reflect the complexity of the phenomenon of sexting. There is, thus, a gap between law and practice.

The Child Sexual Abuse Directive[439] allows Member States to decide whether the exchange of intimate content, including CSAM, among children who have reached the age of consent is criminalised or not.

> *"It shall be within the discretion of Member States to decide whether Article 3(2) and (4) apply to consensual sexual activities between peers, who are close in age and degree of psychological and physical development or maturity, in so far as the acts did not involve any abuse."* (Article 8)

The lack of criminalisation is limited to acts that *"did not involve any abuse or exploitation and no money or other form of remuneration or consideration is given as payment"*.

Only a few Member States, such as Austria, Denmark, Cyprus, and Germany, have made use of the exemption for consensual sexting among children. Most Member States still criminalise sexting, even if they do not prosecute children in practice. Such divergence between theory and practice creates legal uncertainty.[440]

Sexting always involves risks of the intimate content being shared to other peers or even distributed for profit online. However, an abstinence-based response has proven to be ineffective to tackle the issue. Criminalising the consensual exchange of intimate content and material among children is problematic as it stigmatises children and can be a barrier for them to seek help in dealing with online risks and OSEC. It effectively makes children more vulnerable, by preventing them from reporting OSEC to practitioners or competent authorities who would be required to report to law enforcement.[442]

The legislation, including the Child Sexual Abuse Directive, makes a distinction between children who have, and have not, reached the age of consent in order to define a threshold for decriminalising sexting. Using the age of consent can be problematic, particularly for countries with a high age of consent. Worldwide, most countries have set the age of consent at between 16 to 18 years. Within the EU, the age of consent is on average more commonly between 14 to 16 years old.[443] However, reaching the age of consent does

[437] Mishna, F., Milne, E., Cook, C., Slane, A. and Ringrose, J. (2021) 'Unsolicited Sexts and Unwanted Requests for Sexts: Reflecting on the Online Sexual Harassment of Youth', ***Youth & Society.*** **https://doi.org/10.1177/0044118X211058226**; Hartikainen, H., Razi, A. and Wisniewski, P. (2021) 'Safe Sexting: The Advice and Support Adolescents Receive from Peers Regarding Online Sexual Risks', ***Proceedings of the ACM on Human-Computer Interaction***, 5(CSCW1), 42, pp.1-31 **https://doi.org/10.1145/3449116.**

[438] Mishna, F., Milne, E., Cook, C., Slane, A. and Ringrose, J. (2021) 'Unsolicited Sexts and Unwanted Requests for Sexts: Reflecting on the Online Sexual Harassment of Youth'.

[439] Directive (EU) 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA (2011) ***Official Journal of the European Union L 335.*** Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02011L0093-20111217** (Accessed 21 April 2022).

[440] Chatzinikolaou, A. and Lievens, E. (2021) 'Sexting amongst Children and Teenagers : Towards a Policy That Balances Protection and Autonomy', ANSER : ***Five Years of Global Academic Collaboration Building Evidence for Sexual and Reproductive Health and Rights Policies***, pp.82-85. Available at **http://hdl.handle.net/1854/LU-8704128** (Accessed 11 May); Council of Europe Lanzarote Committee (2022) ***Implementation report – The Protection of Children Against Sexual Exploitation and Sexual Abuse Facilitated by Information and Communication Technologies (ICTs) – Addressing the challenges raised by child self-generated sexual images and/or videos***. Available at: **https://rm.coe.int/implementation-report-on-the-2nd-monitoring-round-the-protection-of-ch/1680a619c4** (Accessed 11 May).

[441] Patchin, J. W. and Hinduja, S. (2020) 'It is time to teach safe texting', ***Journal of Adolescent Health,*** 66(2), pp.140-143. **https://doi.org/10.1016/j.jadohealth.2019.10.010**.

[442] Witting, S.K. (2020) ***Child sexual abuse in the digital era : Rethinking legal frameworks and transnational law enforcement collaboration.*** PhD thesis. Leiden University. Available at: **http://hdl.handle.net/1887/96242** (Accessed 11 May).

[443] World Population Review (2022) ***Age of Consent by Country 2022***. Available at: **https://worldpopulationreview.com/country-rankings/age-of-consent-by-country** (Accessed 11 May); SRHR Africa Trust, TrustLaw, Arnold & Porter Kaye Scholer LLP (2018) ***Age of Consent: Global Legal Review***, p.12. Available at: **https://www.satregional.org/wp-content/uploads/2018/05/Age-of-consent-Global-Legal-Review.pdf** (Accessed 11 May).

not translate into sexual autonomy for all children, considering that each child evolves at their own pace. In addition, the age of consent does not reflect the vulnerability of children in situations of trust relationships, authority dynamics, age differences, and how the influence of peers or a partner plays out in their capacity to give consent voluntarily.[444] Therefore, the legislation should not set a specific age for the decriminalisation of consensual sexting, since the vulnerability of a child depends on the individual rather than on their age.[445]

Non-consensual sexting occurs when the child is pressured or coerced into participating in the sexual exchange, or when there is no consent for the sharing or further distribution of intimate content that was initially shared with consent. Sexting can foster harmful behaviour and place teenagers at online and offline risks.[446] Criminal responses to non-consensual sexting can result in exposure to the criminal justice system and have a harmful effect on children. Yet, intervention is needed to address illegal behaviour as full decriminalisation could encourage non-consensual sexting and sharing of harmful content.[447] Prosecution should occur in cases of OSEC.

Sexting, both consensual and non-consensual, can also lead to the creation of **self-generated CSAM**, where children produce sexually explicit content themselves. About three quarters (72%) of webpages actioned during 2021 by Internet Watch Foundation (IWF) were assessed as containing self-generated material. The material mostly depicted girls aged 11 to 13 years old in their bedrooms or another room at home.[448] Children should not end up with a criminal record relating to self-generated material, especially when pressured or coerced into producing such content.[449] Therefore, in line with the Lanzarote Committee recommendations, the EU legislation should clarify that the production and possession of self-generated sexually explicit content shared among consenting children should not lead to prosecution in cases where the children consented to the production of the material, are above the age of consent, and the material does not involve abuse.[450]

4.2.2 EU to require platforms to have effective age verification systems

Recommendation 2: Age, development, and maturity affect the risks children face and how they perceive and respond to them. Most social media applications resort to self-declaration of age which, while cheap and easily implemented, cannot be considered an effective age assurance mechanism. Children under the age of 13 are likely to lie about their age, knowing that otherwise they would be excluded from accessing the service. The EU should broaden the scope, and strengthen the requirements for age verification systems in their Directive to ensure that children have an age-appropriate experience.

CALL FOR ACTION

1. **Require the use of age verification methods for all platforms**. The AVMSD suggests that age verification be in place for video-sharing platforms. This should be extended to all social media platforms.
2. **Add requirements for effective, privacy-preserving age assurance.** Since the Directive currently in place solely requires the age verification tool to be in place, but not to be effective per se, the EU should revise the Directive to incorporate additional requirements that will make age verification stronger and less easy to circumvent. This can include age assurance technologies that verify the stated age of a user through additional data.

[444] Witting, S.K. (2020) ***Child sexual abuse in the digital era: Rethinking legal frameworks and transnational law enforcement collaboration***.

[445] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[446] Tariq, M. U., Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2019) 'A Review of the Gaps and Opportunities of Nudity and Skin Detection Algorithmic Research for the Purpose of Combatting Adolescent Sexting Behaviors'. ***21st International Conference on Human-Computer Interaction***. Orlando, United States, 26-31 July. Springer, Cham. https://doi.org/10.1007/978-3-030-22636-7_6.

[447] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[448] Internet Watch Foundation (2022) ***Annual Report 2021 – Self-generated child sexual abuse***. Available at: https://annualreport2021.iwf.org.uk/trends/selfgenerated. (Accessed 11 May 2022).

[449] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[450] Council of Europe Lanzarote Committee (2022) ***Implementation report – The Protection of Children Against Sexual Exploitation and Sexual Abuse Facilitated by Information and Communication Technologies (ICTs).***

3. **Explore possibilities to help online platforms verify age and identity.** Some countries have co-operation between the government and other industries to verify identities and age. The EU could explore possibilities to see how it could help the age verification process.

RATIONALE

The AVMSD already includes some requirements for online safety design to shield children from content that is not age appropriate and possibly damaging for their wellbeing and development. Age verification is mentioned in the Directive as one of the suggested measures. The AVMSD requires, however, only for the age verification to be in place and not for it to be effective. Most social media platforms use **self-declaration** as age verification, where stating a date of birth suffices. This report has shown that this type of age verification is **easily circumvented**. Children want to be active on certain platforms, including when they are under the age of 13. Simply by entering a different date of birth from their own, they are able to use the platform. This means that many children can easily gain access to content on platforms that is not age appropriate.

Many children in the conducted focus groups mentioned stronger age verification as a solution to better protect them online. As a recommendation to the industry, it was stated that **multiple sources** should be used to strengthen age verification both upon registration and later. This could be done using third-person verification, public databases, as well as age assurance technologies that build a profile of a person based on online habits, contacts, and language use.

As of now, Member States do not have national laws in place that guide platform owners to strengthen age verification. In order to ensure the effectiveness of age verification, it is important to stress the need for a **unified and harmonised approach** across EU Member States (and worldwide), which is currently not the case.[451] The EU should **revise the Directive to require stronger age verification methods and provide minimum standards** for platforms to incorporate. This is important, since platforms can use age verification methods to exclude certain children that do have the right to access platforms and to obtain additional information from users. It is therefore also vital that the EU establishes rules for age verification to be privacy-preserving.[452] Additionally, since the AVMSD only focuses on video-sharing platforms, the EU needs to broaden the scope of the Directive to include all platforms that children want to access.

4.2.3 EU to require all platforms to have transparent reporting mechanisms and referral systems for children reporting OSEC

Recommendation 3: Children in the focus groups highlighted the need for child-friendly and age-appropriate explanations of platforms' reporting mechanisms, and simpler, more visible options for reporting potential violations. The EU should put efforts into making the reporting rate of OSEC and CSAM higher by clarifying what transparency means and providing additional requirements to properly handle reports. The EU should also establish a mandatory referral and report mechanism when a child reports OSEC.

CALL FOR ACTION

1. **Clarify and strengthen the transparency requirements of online reporting in the Directive.** The EU should provide more details about what transparency means to ensure its applicability to all platforms. This also includes setting a time limit for platforms to inform the person making the report about the progress or outcome of their report.
2. **Require platforms to refer children to appropriate help services after they make an online report.** The EU should provide rules that require platforms to refer children who report OSEC to the right help or support service.

[451] Caglar, C. and Nair, A. (2021) *EU Member State Legal Framework.* Available at: https://euconsent.eu/download/eu-member-state-legal-framework/ (Accessed 22 April 2022).

[452] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

3. **Establish mandatory reporting mechanisms in the event that a child reports a form of OSEC.** Platforms should be required to appoint well-being officers (with a clinical background) who would fall under a mandatory child abuse reporting obligation, similar to teachers.

RATIONALE

As part of the safety measures, the AVMSD requires video-sharing platforms to have transparent and user-friendly reporting mechanisms. The Directive also suggests having systems in place where users who reported an issue will get updates about the outcome of their report. The participating children in the focus groups, however, indicated that they are hardly aware of reporting mechanisms and are also unaware of the process after they file a report. Some children complained that they never heard back after reporting. Moreover, **under-reporting** is a known problem when it comes to OSEC. This all illustrates that the measures put in place by the EU are not sufficient.

Therefore, the EU should strengthen the transparency requirement by first of all **clarifying what transparency means.** It should provide minimum legal requirements and guidelines that will help platforms to implement reporting mechanisms that will be more effective. This might include **differentiating** the reporting process for adults and children. If a platform identifies that the user filing a report is a child, different measures should be taken. Platforms should be provided with requirements to enhance their transparency about content and reporting towards their users. This could, for example be setting a **time limit** for receiving an outcome or getting updates after reporting a case.[453]

Additionally, the EU should add elements to the Directive that require platforms to **refer children who file a report to the appropriate services**. This could, for instance be an organisation that deals with sexual exploitation of children. In this way, they will encourage children to take action outside of the platform. If the child decides to contact the referred organisation, they are also likely to receive more targeted information or specialised help. This will enhance their safety.

In addition, when a child reports a form of OSEC, platforms should be required to report these cases to the competent authorities. This could take the form of appointing well-being officers (with a clinical background) who would fall under a **mandatory child abuse reporting obligation** to the extent of the law, similar to when teachers, counsellors, and medical professionals become aware of imminent risks posed to a child. Such safeguards should have both mandatory components and guidelines for protecting young people, and should be applied universally across all companies that digitally engage with children.

The children in the focus groups also called for a stronger reporting system. They recommended making the reporting option **more visible**, such as having a clearly visible danger button. They also recommended providing children with **more options to choose from** when reporting. Just looking at the menu of options to report could already teach children what is appropriate and what is not. The EU could require platforms to incorporate the recommendations they made. The focus group participants also suggested that platforms should preserve reported evidence and directly **send a copy of the report to the police**. Lastly, they also suggested requiring **sufficient staffing** in complaints and reporting departments in order to properly handle reported cases within a reasonable time.

4.2.4 EU to make online platforms legally accountable for establishing minimum safety standards to keep children safe

Recommendation 4: Platforms should be held accountable for establishing minimum safety features to safeguard the children that use their platform. The goal of this legislation would be to require online platforms to adopt minimum standards for safeguarding children through the deployment of safety by design measures and carrying out regular Child Rights Impact Assessments.

[453] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

CALL FOR ACTION

- **The EU should make platforms accountable for establishing minimum safety standards**. The policies should be consistent for all service providers, with enforceable minimum standards and a code of practice for safety by design that covers all elements of a service, its features, and functionalities, with full oversight provided by a well-resourced regulator.
- **The minimum standards should include requirements for child rights impact assessments (CRIAs)**. Platforms should regularly publish child rights impact assessments of their services, which should be monitored by the EU.

RATIONALE

As demonstrated in this research, children are not safe online. Platforms have adopted some measures to protect children online, yet they are largely insufficient. In fact, some of the design and algorithms used by platforms contribute to putting children at risk. Platforms are designed to facilitate content sharing by nudging users towards sharing widely and publicly. Since children underestimate risks online and prioritise the social benefits of oversharing, design that facilitates over-disclosure leads to higher risks for children. Platforms should be held accountable for design that put children at risk.

When it comes to child users, platforms should adopt some **minimum safety features** that are shown to be effective in minimising risks. For instance, this can take the form of requiring that platforms use privacy by default settings for children's accounts, automatically setting their accounts to private. As we know that the status quo bias leads users not to change their default settings, such a change would effectively strengthen the safety of child users (see also the recommendation for the industry to deploy intelligent privacy features customised to keep all children safe). Safety features should also be implemented on the devices (iOS and Android) of child users.[454] The industry has already started to make changes to this end. The EU could help **codify such minimum safety standards** and hold platforms accountable to implement them.

In line with General comment No. 25 of the United Nations Committee on the Rights of the Child (UN CRC),[455] the EU should require that platforms conduct **child rights impact assessments** of their services and operations, including on the safety implications of their services designed for, or accessed by child users. Such impact assessments should be publicly available and closely monitored by the EU.

Overall, the EU should aim for the industry to adhere to the highest standards of safety for children and hold companies accountable for ensuring safety, taking into account the best interests of children.

4.2.5 EU to support and strengthen initiatives aimed at education, awareness, action research, and offender interventions

Recommendation 5: Education is an important measure for preventing and tackling OSEC. Instead of shielding children from risks altogether, it recognises children's need to be online and teaches them how to recognise and deal with risks, thereby building digital resilience. As the EU has a supporting competence with Member States for education, the EU could play a significant role in funding key initiatives that promote digital skills and literacy, spread awareness, educate on consent online and decrease online offending.

[454] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[455] UN CRC (2021) **General comment No. 25 (2021) on children's rights in relation to the digital environment**. Available at: https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation (Accessed: 27 April 2022).

CALL FOR ACTION

1. **Initiate and fund education and awareness programs.** The EU should support Member States by offering or funding programmes that promote healthy (online) sexual behaviour or that promote awareness about OSEC and CSAM. These programmes should include a broad range of skills that children can deploy, either offline or online, to build their (digital) resilience.
2. **Fund safe online support platforms.** The EU could fund online platforms where adolescents can discuss sensitive topics that they might feel uncomfortable sharing with adults. The EU could support these initiatives by funding the securing of these platforms so they offer a safe environment for children, or by having professionals present on the platforms that could monitor conversations and provide children with professional advice.
3. **Fund action research on OSEC and CSAM.** More research should be funded by the EU that specifically focuses on effective safety design, prevention, and intervention. The EU could also support in disseminating evidence provided by these research reports and making sure that the knowledge is taken up by the industry.
4. **Fund interventions aimed at OSEC offenders.** The EU should put more effort into identifying effective interventions for OSEC offenders. They should increase their funding of these interventions, focusing on the demand side of OSEC to eventually also protect children further.

RATIONALE

EDUCATION AND AWARENESS PROGRAMMES

This report showed that an abstinence-based approach, where children are shielded from risks of OSEC altogether, is not always feasible or healthy for developing children. Instead, children should be educated about how to navigate the digital world.[456] Equipping children with the right knowledge and skills can serve as a protective factor against online risks.[457] Ali, Haykal and Youssef (2021) call this **'building digital resilience'.**[458] The children in the focus groups also asked for more awareness-raising around the topics. They suggested that television or public figures, such as influencers, should be used to spread the message in an attractive way. The EU could support these campaigns through funding.

The education on this topic should cover **different areas**. Firstly, children should have knowledge about privacy, online risks and risk management. This includes teaching about how sexual exploiters approach young people and what the red flags are.[459] It is also relevant to teach children about self-awareness online, the visibility of their online information, privacy, and thus in general what the consequence of their internet use is.[460] Important to note is that education about online safety should not only focus on the negatives, instilling fear on children. Instead, it is important to empower children by positively teaching children about healthy sexual development and relationships. This should include knowledge about sexual identity, emotions, dynamics of a healthy relationship, and sexual and reproductive health issues. The EU could fund initiatives by organisations that promote these positive behaviours, instead of focusing on the law enforcement side.

In order to effectively change the online behaviour of children, education should not only be focused on giving information about risks, but also teaching certain **skills**.[461] Various sources have suggested areas that could be covered by this type of skills education for children, such as credibility assessments, personal security,

[456] Badillo-Urquiola, K., Razi, A., Edwards, J. and Wisniewski, P. (2020) 'Children's perspectives on human sex trafficking prevention education', *ACM International Conference on Supporting Group Work (GROUP '20)*, Sanibel Island, United States, 6-8 January. ACM, New York, United States, pp.123-126. https://doi.org/10.1145/3323994.3369889.

[457] Livingstone, S. and Smith, P. K. (2014) 'Annual Research Review: Harms experienced by child users of online and mobile technologies: the nature, prevalence and management of sexual and aggressive risks in the digital age', *The Journal of Child Psychology and Psychiatry,* 55(6), pp.635-54. https://doi.org/10.1111/jcpp.12197.

[458] Livingstone, S. and Mason, J. (2015) *Sexual rights and sexual risks among youth online*. Available at: http://eprints.lse.ac.uk/id/eprint/64567 (Accessed 11 May).

[459] Finkelhor, D., Jones, L. and Mitchell, K. (2021) 'Teaching Privacy: A flawed strategy for children's online safety', *Child Abuse & Neglect*, 117. https://doi.org/10.1016/j.chiabu.2021.105064.

[460] Smith, P. K. and Livingstone, S. (2017) 'Child Users of Online and Mobile Technologies – Risks, Harms and Intervention', *Child Psychology and Psychiatry: Frameworks for Clinical Training and Practice*, pp.141-148. https://doi.org/10.1002/9781119170235.ch17.

[461] Quayle, E. and Koukopoulus, N. (2018) 'Deterrence of Online Child Sexual Abuse and Exploitation', *Policing*, p.13(3), pp.345-362. https://doi.org/10.1093/police/pay028.

device security,[462] information navigation, and communication.[463] There are many skills that are helpful for preventing offline and online victimisation and offending. The risks associated with either offline or online abuse often stem from the **same roots or follow the same patterns**.[464] Therefore, it is important to teach children skills such as critical thinking, social skills,[465] and empathy in general. Equipping children with this elaborate knowledge and skill set could contribute to more self-awareness and interpersonal skills that will lessen their risk for offending or being victimised.[466]

Online safety education could be connected to offline safety training. Education about offline safety has been around longer than online safety education. Therefore, the latter could draw on the lessons from offline safety to enhance their programme development. Sometimes making the comparison between offline and online risks helps to make the teaching more **tangible** and thus real. It could also be easier to incorporate online risks in existing offline education programmes instead of having to design a separate curriculum.[467] The EU could therefore further fund existing overall health and life skill programmes, where they could push for a link between the offline and online world. This could be, for instance, a mental health or sexual reproductive health programme. Another idea to include is the use of OSEC survivors in these education programmes to provide real-world examples of online risks and the impact OSEC can have on someone's life.

It is also worth teaching children about **what to do after something has happened to them**, whether offline or online. Keeping evidence, how to report someone, through which channels, and what will happen after reporting are all relevant for children to know, and will perhaps make them feel more confident that if something happens, they will know what to do. If children know what kind of evidence is needed for the police to properly handle a case, this could increase their willingness to report and the strength of their report.[468]

Education and awareness programmes are not limited to children. **Platform owners** should also be educated about online risks and what to focus on. Additionally, in various research, parents have expressed feeling helpless and lost, as they are less familiar with social media for instance than their children.[469] This results in them not feeling equipped to teach their children the necessary skills.[470] Therefore, it is crucial that both professionals working with children and parents have sufficient tools and knowledge to teach online safety to children in an age-appropriate manner.[471] Age-appropriate means in a way that is understandable to children, but also in a manner that corresponds with their developmental stage, giving them the right amount of agency and opportunities to explore and learn from risks.[472]

Educating professionals and parents about online risks should also include the **signs children display when they are harmed and what they, as an adult can do when a child signals this**.[473] This could involve having information on the harmful consequences, but also the legal landscape surrounding offline and online child

[462] Stoilova, M., Livingstone, S. and Khazbak, R. (2021) *Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children's internet use and outcomes*. Available at: **https://www.unicef-irc.org/publications/pdf/Investigating-Risks-and-Opportunities-for-Children-in-a-Digital-World.pdf** (Accessed 27 April 2022).

[463] ibid

[464] Finkelhor, D., Walsh, K., Jones, L., Mitchell, K. and Collier, A. (2021) 'Youth Internet Safety Education: Aligning Programs with the Evidence Base', **Trauma, Violence and Abuse**, 22(5), pp.1233-1247. **https://doi.org/10.1177/1524838020916257**.

[465] Stoilova, M., Livingstone, S. and Khazbak, R. (2021) *Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children's internet use and outcomes*.

[466] Livingstone, S. and Mason, J. (2015) *Sexual rights and sexual risks among youth online*.

[467] Finkelhor, D., Jones, L. and Mitchell, K. (2021) 'Teaching Privacy: A flawed strategy for children's online safety'.

[468] Smith, P. K. and Livingstone, S. (2017) 'Child Users of Online and Mobile Technologies – Risks, Harms and Intervention'.

[469] Livingstone, S., Blum-Ross, A., Pavlick, J. and Ólafsson, K. (2018) *In the digital home, how do parents support their children and who supports them?* Available at: **https://www.lse.ac.uk/media-and-communications/assets/documents/research/preparing-for-a-digital-future/P4DF-Survey-Report-1-In-the-digital-home.pdf** (Accessed 25 April 2022).

[470] Erickson, L. B., Wisniewski, P., Xu, H., Carroll, J. M., Rosson, M. B. and Perkins, D. F. (2016) 'The boundaries between: Parental involvement in a teen's online world', *Journal of the Association for Information Science and Technology,* 67(6), pp.1384-1403. **https://doi.org/10.1002/asi.23450**.

[471] Livingstone, S., Mascheroni, G., Dreier, M., Chaudron, S. and Lage, K. (2015) *How parents of young children manage digital devices at home: the role of income, education and parental style.* Available at: **http://eprints.lse.ac.uk/id/eprint/63378** (Accessed 11 May).

[472] Livingstone, S. and Mason, J. (2015) *Sexual rights and sexual risks among youth online*.

[473] Smith, P. K. and Livingstone, S. (2017) 'Child Users of Online and Mobile Technologies – Risks, Harms and Intervention'.

abuse.[474] When a child confides in an adult after victimisation, the attitude of the adult is also very important. A **supportive environment** can be a protective factor to shield children from online risks, as the children might find it easier to consult adults when in distress. Adults can then quickly jump in to assist and help mitigate the negative consequences of facing online adversity.[475]

Parents have an extra important role, as they are also in charge of **enabling or restricting** their children from using the internet. Parents should be taught that restricting their children's internet use is not completely harmless. Less internet use might lead to less exposure to risk, but also less enjoyment of the benefits of the internet. Moreover, when parents restrict their children too much, they also discourage children from involving their parents in their internet use. Parents that enable their children to use the internet are positively associated with **agency** and their children **asking for their guidance**.[476]

SAFE ONLINE SUPPORT PLATFORMS

Adults striving to keep children safe online is of course very important, but peer support also proves to be very valuable to children. In the focus group discussions, a little over half of the children indicated that they feel more comfortable talking to a friend than to a parent when encountering something distressing online. Children feel that adults do not understand their problems and might punish them for getting in trouble. Friends, however, can more easily relate due similar experiences online, the use of similar platforms, and their age.

> *"Sometimes, friends can give better advice than parents. This is because parents do not understand our lifestyles, how we live our lives. It is more pleasant or relaxed to talk with friends than parents."*
> (Focus group discussion, child, Thailand)

Children should therefore be given a **safe space** where they can seek and receive this peer support. Such platforms enable children to discuss what can be uncomfortable or awkward topics that involve sexuality and online harm. Some platforms are anonymous, making it easier for children to discuss these sensitive issues. A downside of anonymity is that it could also lower the threshold for children to bully others, as their identity is unknown. The EU could help to create these platforms and also ensure safety. It could also fund initiatives by existing helplines to establish a peer-support space.[477]

The EU could also ensure **funding for professionals** to be part of these platforms. Firstly, this could be professionals who moderate the platform to make sure it is a safe place. Secondly, this could be professionals trained on these topics, who could be part of the platform and provide children with advice. This might lower the threshold for children to reach out for help on the platform. Such professionals could also refer children to real therapists when they feel this is necessary.

ACTION RESEARCH ON OSEC AND CSAM

A clear recommendation that stems from the literature is that there is a need for more funding from the EU for research. There is currently a lot of research into the phenomena OSEC and CSAM that often does not result in practically usable design suggestions. The EU should make funds available for **action research** that specifically focuses identifying evidence-based interventions and safety designs.

OFFENDER INTERVENTIONS

A last area of funding is aimed at child or adult OSEC offenders. These interventions can be deployed at different stages: prevention, at the time of the event, and after offending. For prevention, counselling and self-help courses could help.[478] There are existing programmes for people who experience **inappropriate thoughts** where they can be helping to deal with those thoughts.[479] **Broader awareness and education** might not

[474] Patchin, J. W. and Hinduja, S. (2018) 'Sextortion Among Adolescents: Results From a National Survey of U.S. Youth', *Sexual Abuse*, 32(1), pp.30-51. https://doi.org/10.1177/1079063218800469.

[475] Ali, S., Haykal, H. A., Youssef, E. Y. M. (2021) 'Child Sexual Abuse and the Internet- A systematic review', *Human Arenas*, 4(1). https://doi.org/10.1007/s42087-021-00228-9.

[476] Livingstone, S., Ólafsson, K., Helsper, E. J., Lupianez-Villanueva, F., Veltri, G. A. and Folkvord, F. (2017) 'Maximizing opportunities and minimizing risks for children online: the role of digital skills in emerging strategies of parental mediation', *Journal of Communication*, 67(1), pp.82-105. https://doi.org/10.1111/jcom.12277.

[477] Feedback from senior experts/participants in the online workshop conducted as part of the present research.

[478] Quayle, E. and Koukopoulus, N. (2018) 'Deterrence of Online Child Sexual Abuse and Exploitation'.

[479] Gillespie, S. M., Bailey, A., Squire, T., Carey, M. L., Eldridge, H. J. and Beech, A. R. (2018) 'An Evaluation of A Community-Based Psycho-Educational Program for Users of Child Sexual Exploitation Material', *Sexual Abuse*, 30(2), pp.169-191. https://doi.org/10.1177/1079063216639591.

directly stop offending, but it could prompt offenders to seek help or could educate potential victims further, thereby reducing their availability and willingness to engage.[480] In order for prevention to be effective, just before someone is going to offend, the risks should be increased and the benefits reduced, making it less attractive for offenders to give into their needs.[481] To increase the risks, **intervention tactics** could be used where warning messages pop up about the risks of a certain activity.[482]

Lastly, there are also possibilities for education after the offence has already happened. In their research, Gillespie et al. (2018) looked at a '**community-based psycho-educational group-work program**'[483] in the United Kingdom that targets adult men who view child sexual exploitation material. The participants follow ten group sessions where they discuss and work on different topics. This programmes thus heavily depends on education, for instance on openly discussing sexual fantasies, but also teaching about social skills and empathy. Their study showed improvements in the following areas: less depression, less anxiety, better social skills and increased self-esteem. Their research shows that these types of programmes can have an effect on difficulties that offenders face.[484] The areas that this programme focused on are also often present in other interventions aimed at perpetrators of OSEC. It should, however, be noted that most of the interventions are focused on people who have already committed an offence, thus aiming at preventing reoffending instead of offending in the first place or desisting from engaging with OSEC altogether.[485]

[480] Quayle, E. and Koukopoulus, N. (2018) 'Deterrence of Online Child Sexual Abuse and Exploitation'.

[481] ibid

[482] ibid

[483] Gillespie, S. M., Bailey, A., Squire, T., Carey, M. L., Eldridge, H. J. and Beech, A. R. (2018) 'An Evaluation of A Community-Based Psycho-Educational Program for Users of Child Sexual Exploitation Material. Sexual Abuse', p.173.

[484] Gillespie, S. M., Bailey, A., Squire, T., Carey, M. L., Eldridge, H. J. and Beech, A. R. (2018) 'An Evaluation of A Community-Based Psycho-Educational Program for Users of Child Sexual Exploitation Material. Sexual Abuse.

[485] Perkins, D., Merdian, H., Schumacher, B., Bradshaw, H. and Stevanovic, J. (2018) *Interventions for perpetrators of online child sexual exploitation – A scoping review and gap analysis.* Available at: https://www.csacentre.org.uk/documents/online-cse-interventions/ (Accessed 11 May).

5. Conclusions

Increasing numbers of children, and increasingly young children, are using the internet. Over the years, the opportunities the internet offers have increased, but with that, the risks have also increased. One of those risks is online sexual exploitation of children (OSEC), with a peak in its prevalence during the COVID-pandemic. OSEC is a form of gender-based violence that can have a detrimental impact on children's wellbeing and development. It is thus crucial to prevent children from being victimised. Online platforms and their design could play a significant role in keeping children safe online, but the measures currently in place are not sufficient. EU policies are also not stringent enough to force online platforms to change their designs. To contribute to better understanding of child safety by design, this research used in-depth desk research, focus group discussions with children in 10 countries, and a panel of international experts in online safety. As a result, this research identified design solutions and EU policy recommendations to better protect children against OSEC. First, the most important findings of the report will be summarised. In section [x], the recommendations for both the industry and the EU will be presented and explained.

5.1 CHILD SAFETY BY DESIGN IS CHALLENGING DUE TO THE MANY FACTORS ASSOCIATED WITH OSEC AND CONFLICTING VALUES

This report found that safety by design is both complex, in terms of the number of associating and intersecting factors that need to be considered for design to be effective, and complicated, in terms of the difficulty in addressing those diverse factors. Gender, age, disability, and sexual orientation all play a role, and their impact can differ per child as each personality is different. Children who have a disability or intellectual impairment are at risk, as well as children who identify as LGBTQI+ or are in doubt about their sexuality. As OSEC is a form of gender-based violence, gender influences the risk of OSEC tremendously. Unequal gender norms, with male dominance and female submissiveness, results in girls being more often victimised than boys and offenders usually being men. There are also some specific risks based on gender. Girls are more likely than boys to be victims of OSEC, to be depicted in child sexual abuse material (CSAM), and to be blamed for their victimisation; boys are more likely to be victims of sexual abuse outside the family or be exposed to pornography. Boys are also less likely to report their victimisation than girls.

Another big factor is age. At the onset of puberty, children are most vulnerable and increasingly younger children are at risk online. In the pre-teen and teenage years, children go through many changes, both biologically and cognitively, leading them to engage in more risky and sexual activities online, while not having the full cognitive capacity to perceive all the risks. Choosing appropriate design for children is complicated as they quickly develop and their needs change as they evolve. Even within one age group, children are a diverse group with varying levels of maturity and vulnerabilities. On the one hand, children should be kept safe, as the consequences of OSEC can be severe and children are not fully able to adequately deal with risks. On the other hand, children should increasingly be given more responsibility and agency in navigating the online world, as it offers them many advantages. Design should be able to reflect this cognitive development with increasing possibilities.

Designers of online platforms do not have an easy task, as they should take this development into account, tailoring their design to the needs of different age groups, as well as accounting for gender issues, vulnerabilities, and conflicting values such as freedom of expression, privacy, building resilience, and protection. The platforms also find themselves in a tricky situation, where they have to sustain a business where more users and more interaction between users is more profitable than making their platform more restrictive.

Because of these diverse needs and conflicting values, there is no silver bullet solution. Instead, tackling OSEC requires a set of solutions tailored to various age groups and risks. Child safety by design is thus a very complex and complicated field.

5.2 KNOWLEDGE IS A POWERFUL PREVENTIVE MEASURE

Being able to effectively tackle OSEC and its risks starts with understanding the topic fully. Therefore, the deeper understanding of OSEC and child safety by design provided by this report proves to be important not only as a contribution to general knowledge, but also as a prevention tool. Knowledge about the patterns and risks of OSEC could in itself also contribute to more caution for multiple stakeholders:

1. For **designers** of online platforms so they know where their focus of safety design should lie.
2. For **children** so they can protect themselves better, recognise red flags, and know what inappropriate behaviour is.
3. For **adults** (parents and professionals) so they know what to teach their children and which signals they could spot when children face online harm.
4. For **policy-makers** so they adopt laws that can effectively protect children online.

This report uncovers various vulnerabilities that lead to an increased risk of OSEC. Sexual predators victimising children are likely to be calculated individuals, assessing the vulnerability of children to decide who to target. They can also take advantage of these vulnerabilities and use companionship, manipulation, and flattery to get closer to the child. It has become easier and more acceptable for children to be friends with strangers, increasing the risks of OSEC. Additionally, children are more likely to be victimised by people they know. Children have expressed having more difficulty shutting down conversations with people they know, making this a dangerous dynamic. **Children themselves can also be part of the problem**, as they can commit OSEC and view or generate CSAM. As explained previously, children in puberty increasingly explore their sexuality, with a peak in early adolescence. The internet provides many possibilities for this. Knowing these patterns and dynamics can help the afore-mentioned stakeholders to detect patterns of OSEC earlier and to intervene at a stage where harm can still be prevented.

Children are also part of the solution as it is more effective to involve them in assessing risks online and guiding them in dealing with the risks they face, thereby empowering them to protect themselves. Teaching children digital literacy should encompass different kinds of information and various skills. This education could be incorporated in existing lessons about real-world risks, as the roots and patterns are often similar, and offline safety training is already more established. In addition to information about online risks, children should be taught general life skills such as communication, empathy, boundaries, and consent, as well as skills that will help them online, such as assessing the credibility of information and learning about their digital footprint. Professionals and parents should be equipped with tools to further support this kind of education.

For parents, this also includes knowledge on how to help their children navigate the online world. Parental monitoring and restriction can be used for younger children up to the age of 12, whereas communication and self-regulation should be promoted for children over this age. It is important for children to build digital resilience and learn by experience how to prevent and deal with online risks. Shielding children too much from adversity could also hamper their development. An overly restrictive approach can also damage the parent-child relationship and increases the risk of children not telling their parents about distressing encounters online.

Knowledge on the topic is thus very important, but not all parts of online safety child safety by design are well-researched. This research uncovered various areas that need additional attention. Firstly, this involves action research that looks into the effectiveness of online tools and designs. There is a fair amount of research focusing on the online risks themselves, but not how to effectively tackle those. Research that will contribute to evidence based safety designs is necessary. Additionally, there is not enough research on how vulnerabilities, such as sexual orientation, gender and age, play out in the online world. Little research was found, for instance, on keeping younger children safe online. With regard to gender, there are many estimates that girls are more likely to be groomed, for instance. The exact influence of gender and age on risk is much less researched. A better understanding of these dimensions can contribute to better preventative and support approaches. Lastly, this research found that children can also be part of the problem of online risk. There have not been many studies that focus specifically on the risks and pathways to child offending and the duality of being a victim and being an offender. More research on these topics can better target preventative interventions as well.

5.3 MEASURES CURRENTLY IN PLACE ARE NOT SUFFICIENT TO KEEP CHILDREN SAFE

Currently, online platforms have different measures in place to support child protection. This includes general measures such as age verification, reporting, and community guidelines. Some platforms provide special privacy settings for children by default, such as having a private profile. Other platforms already have some monitoring designs, where they use classifiers or artificial intelligence (AI) to detect risks and warn users when they are about to engage in risky behaviour. Overall, it could be stated that platforms do have safety measures in place, but that these are not currently sufficient. It is a problem, for instance, that the measures taken are not uniform across platforms. Furthermore, safety designs are usually not tailored to gender or different age groups, thus neglecting the diverse needs of children. Existing measures could also be easy to circumvent and are not so stringent that they effectively eliminate OSEC.

For the legal and policy level, the same conclusion applies. Some legislation and requirements for tackling OSEC already exist, but there is still a lot of room for improvement. This report has shown that OSEC is a widespread risk with potentially adverse effects for children. No parent would send their child to play in a physical space with as much likelihood of harm as the online world, yet children spend an increasing amount of time online without adequate safety. Why do we not approach the online world as we do the physical spaces? It is time **to revise the existing EU legal framework** and incorporate more measures and restrictions so that children can benefit from all the possibilities and opportunities the internet offers.

In the next section, the recommendations that stem from the literature, experts and children, will be explained, providing concrete suggestions for improvement.

5.4 RECOMMENDATIONS TO THE EU AND THE INDUSTRY

This research produced several recommendations to strengthen online safety by design in seven areas. An overview of these recommendations are tabled below. The first area is **awareness**. As discussed prior, skills and knowledge about online safety can be a powerful preventive tool. Children together with adults should be educated on the topic and equipped with a skillset that will help them make healthy decisions and recognise risks online. The EU is advised to support and further **strengthen programmes that promote online safety, as well as action research** that aims to identify effective design solutions and interventions. The children in the focus groups also emphasised that there should be more awareness, both in the visibility of rules and regulations on online platforms and through campaigns. The children suggested using popular media, but also popular public figures to spread the message about online safety.

Second, **age verification and assurance** should be strengthened.As of now, the EU just requires the age verification to be in place. The EU should therefore revise its Directive to require all platforms to have **an *effective* age verification system**. Online platforms could improve the effectiveness of their systems by **using multiple sources** to verify the age of a user, such as looking for that person in public databases or using technology that will estimate someone's age by analysing language, friends and habits online. The children in the focus groups also suggested that age verification should be strengthened and proposed the use of official documents.

A third area of recommendations focuses on **specialised features for children** in platform designs. For the EU, this means that they should **make platform holders legally** accountable for the risks that children face on their online platform. To keep children safer in design, the industry should develop features that are tailored to different age groups of their users. For children this means not a one size fits all approach for the design, but incorporating **privacy features by default that increasingly give the child more responsibility and freedom on the platform**. In the focus group discussions, the children also proposed to have different settings for different age groups. Additionally, they proposed that content featuring children should have additional safety guarantees, such as the content not being able to be downloaded.

Then, there should be additional safety features that involve **content and user monitoring**. The industry should implement technologies that will **actively look for and identify risks** on platforms. When identified, there should be an intervention, for instance by **warning messaging or educational prompts**. The children of the focus groups also proposed that platforms should keep them more safe by adding extra pop-up and warning messages that will help them rethink whether it is safe what they are doing.

Fifth, **parents** should also play a role in keeping their children safe online. In the focus groups, the majority of children mentioned that their parents are very important in keeping them safe online. Online platform holders should therefore deploy **parental control features** that work both for parents as for children. This means not overly monitoring and restricting the use of the internet of children, but as earlier mentioned, tailoring the parental control features to the developmental stage of the child.

The sixth area focuses on reporting and referrals. Momentarily, there is a huge underreporting rate of OSEC and CSAM. The children in the focus group discussions also mentioned that if they reported something to the platform, they rarely heard something back. The EU should add requirements to their Directive to **increase the transparency of reporting mechanisms**. Additionally, they should require online platforms to **refer reports and children who made an OSEC related report to appropriate services**. In the focus groups, the children also asked for stronger consequences after violations online, as they felt that the punishments were not sufficient to make violators stop.

Lastly, the EU should take a closer look at the criminalisation of online behaviour and recognize the difference between consensual and non-consensual sharing of intimate content. As a recommendation, the **EU should decriminalise the exchange of this type of content among children** in its revision of the Child Sexual Abuse Directive.

Table 6 - Overview of the recommendations

	EU	Industry	Children recommend
Awareness	Support and further strengthen key initiatives aimed at **education, awareness, action research and offender intervention**	Toy Safety Directive – to be revised	Create more **visible** rules, regulations and reporting measures and **using popular media and persons** to deliver safety messages
Age verification and assurance	Require platforms to have an effective age verification system	Strengthen age verification and assurance by **using multiple sources**	Strengthen age and identity verification through **official documents**
Specialised child safety features	Make online platforms **legally accountable** for for establishing minimum safety standards to keep children safe	Develop features that **encourage child / teen empowerment and self-regulation** and deploy **intelligent privacy** features customised to keep all children safe	**Differentiate** for **different age groups**, have intelligent safety features and to have **extra safety** measures concerning **content** that involves children
Content and user monitoring, risk identification and warning messaging		Implement technologies that **identify risks** combined with risk **mitigation strategies**	Intervene in potentially harmful situations by **pop-ups and warnings**
Parental control		Deploy **parental control features** that work for all children across all age groups	**Parents** need to be **involved** in keeping their children safe
Reporting and removal of inappropriate content and users	Require all platforms to have **transparent reporting mechanisms and referral systems** when children are reporting OSEC		Having **stronger consequences after violations**
Criminalisation of online behaviour	**Decriminalise exchange of intimate content** and material among children in its revision of the Child Sexual Abuse Directive		

Bibliography

Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., Komanduri, S., Leon, P. G., Sadeh, N., Schaub, F., Sleeper, M., Wang, Y. and Wilson, S. (2017) 'Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online', *ACM Computing Surveys,* 50(3), 44, pp.1-41. **https://doi.org/10.1145/3054926**.

Ali, S., Haykal, H. A., Youssef, E. Y. M. (2021) 'Child Sexual Abuse and the Internet- A systematic review', *Human Arenas*, 4(1). **https://doi.org/10.1007/s42087-021-00228-9**.

Anderson, P., Zuo, Z., Yang, L. and Qu, Y. (2019) 'An Intelligent Online Grooming Detection System Using AI Technologies', *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, New Orleans, United States, 23-26 June. IEEE, pp.1-6. **https://doi.org/10.1109/FUZZ-IEEE.2019.8858973**.

Assal, H. and Chiasson, C. (2019) '"Think secure from the beginning": A Survey with Software Developers'. *CHI Conference on Human Factors in Computing Systems Proceedings (CHI '19)*. Glasgow, United Kingdom, 4-9 May. ACM, New York, United States. 13 pages. **https://doi.org/10.1145/3290605.3300519**.

Babchishin, K. M., Seto, M. C., Fazel, S. and Långström, N. (2019) 'Are There Early Risk Markers for Pedophilia? A Nationwide Case-Control Study of Child Sexual Exploitation Material Offenders', *The Journal of Sex Research*, 56(2), pp.203–212. **https://doi.org/10.1080/00224499.2018.1492694**.

Badillo-Urquiola, K., Chouhan, C., Chancellor, S, De Choudhary, M. and Wisniewski, P. (2019) 'Beyond Parental Control: Designing Adolescent Online Safety Apps Using Value Sensitive Design', *Journal of Adolescent Research*, 35(1), pp.147-175. **https://doi.org/10.1177/0743558419884692**.

Badillo-Urquiola, K., Page, X. and Wisniewski, P. (2019) 'Risk vs. Restriction: The Tension between Providing a Sense of Normalcy and Keeping Foster Teens Safe Online', *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, United States, Paper 267, pp.1-14. **https://doi.org/10.1145/3290605.3300497**.

Badillo-Urquiola, K., Razi, A., Edwards, J. and Wisniewski, P. (2020) 'Children's perspectives on human sex trafficking prevention education', *ACM International Conference on Supporting Group Work (GROUP '20)*, Sanibel Island, United States, 6-8 January. ACM, New York, United States, pp.123-126. **https://doi.org/10.1145/3323994.3369889**.

Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E. and Wisniewski, P. (2019) 'Stranger Danger! Social Media App Features Co-designed with Children to Keep Them Safe Online', *18th ACM International Conference on Interaction Design and Children*, Boise, United States, 12-15 June. ACM, New York, United States, pp.394-406. **https://doi.org/10.1145/3311927.3323133**.

Bates, L. (2015). *Everyday Sexism*. Simon & Schuster UK.

Basaran, C., Yoon, H. J., Ra, H. K., Son, S. H., Park, T. and Ko, J. G. (2014) 'Classifying children with 3D depth cameras for enabling children's safety applications', *ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*, Seattle, United States, 13-17 September. ACM, New York, United States, pp. 343-347. **https://doi.org/10.1145/2632048.2636074**.

Belton, E. and Hollis, V. (2016) *A Review of the Research on Children and Young People who Display Harmful Sexual Behaviour Online*. Available at: **https://learning.nspcc.org.uk/media/1198/review-children-young-people-harmful-sexual-behaviour-online.pdf** (Accessed 25 April 2022).

Benavente, B., Ballester Brage, L., Pich Solé, J. and Pereda Beltrán, N. (2021) 'Risk Factors for Commercial Sexual Exploitation of Children and Adolescents: Results of an International Delphi Panel', *Psicothema*, 33(3), pp.449-455. Available at: **http://www.psicothema.com/pdf/4691.pdf** (Accessed 18 April 2022).

Black, P. J., Wollis, M., Woodworth, M. and Hancock, J. T. (2015) 'A linguistic analysis of grooming strategies of online child sex offenders: Implications for our understanding of predatory sexual behavior in an increasingly computer-mediated world'. *Child Abuse & Neglect*, 44, pp.140-149. **https://doi.org/10.1016/j.chiabu.2014.12.004**.

Bracket Foundation (2019) *Artificial Intelligence – Combating Online Sexual Abuse of Children*. Available at: **https://respect.international/wp-content/uploads/2019/11/AI-Combating-online-sexual-abuse-of-children-Bracket-Foundation-2019.pdf** (Accessed 25 April 2022).

Bradbury, P. and Martellozzo, E. (2021) ''Lucky Boy!'; Public Perceptions of Child Sexual Offending Committed by Women', *Journal of Victimology and Victim Justice*. **https://doi.org/10.1177/25166069211060091**.

Burén, J. (2020) *Sexting among adolescents: A gendered online phenomenon, related to individual and social determinants*. PhD thesis. University of Gothenburg.

Burén, J. and Lunde, C. (2018) 'Sexting Among Adolescents: A Nuanced and Gendered Online Challenge for Young People', *Computers in Human Behavior*, 85, pp.210-217. **https://doi.org/10.1016/j.chb.2018.02.003**.

Bursztein, E., Clarke,E., DeLaune, M.,Elifff, D. M., Hsu, N., Olson, N., Shehan, J., Thakur, M., Thomas, K. and Bright, T. (2019) 'Rethinking the Detection of Child Sexual Abuse Imagery on the Internet', *World Wide Web Conference (WWW '19)*. San Francisco, United States, 13-17 May. ACM, New York, United States, pp.2601–2607. **https://doi.org/10.1145/3308558.3313482**.

Cabello-Hutt, T., Cabello, P. and Claro, M. (2018) 'Online opportunities and risks for children and adolescents: The role of digital skills, age, gender and parental mediation in Brazil', *New Media & Society*. 20(7), pp.2411-2431. **https://doi.org/10.1177/1461444817724168**.

Caglar, C. and Nair, A. (2021) *EU Member State Legal Framework*. Available at: **https://euconsent.eu/download/eu-member-state-legal-framework/** (Accessed 22 April 2022).

Canadian Centre for Child Protection Inc., *Survivors' Survey – Executive Summary 2017*. Available at: **https://www.protectchildren.ca/pdfs/C3P_SurvivorsSurveyExecutiveSummary2017_en.pdf** (Accessed 18 April 2022).

Chatzinikolaou, A. and Lievens, E. (2021) 'Sexting amongst Children and Teenagers : Towards a Policy That Balances Protection and Autonomy', *ANSER : Five Years of Global Academic Collaboration Building Evidence for Sexual and Reproductive Health and Rights Policies,* pp.82-85. Available at **http://hdl.handle.net/1854/LU-8704128** (Accessed 11 May).

Cockbain, E., Ashby, M. and Brayley, H. (2015) 'Immaterial boys? A large-scale exploration of gender-based differences in child sexual exploitation service users', *Sexual Abuse*, p.5. **https://doi.org/10.1177/1079063215616817**.

Cockbain, E. and Brayley, H. (2012) 'Child sexual exploitation and youth offending: A research note', *European Journal of Criminology*, 9(6), pp.689-700. **https://doi.org/10.1177/1477370812453401**.

'Communication from the Commission Guidelines on the practical application of the essential functionality criterion of the definition of a 'video-sharing platform service' under the Audiovisual Media Services Directive (2020/C 223/02)' (2020) *Official Journal of the European Union C* 223/3. Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C_.2020.223.01.0003.01.ENG&toc=OJ:C:2020:223:TOC** (Accessed 22 April 2022).

Consolidated version of the Treaty on European Union (TEU) (2016) *Official Journal of the European Union C* 202/1. Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A12016M%2FTXT** (Accessed 29 April 2022).

Consolidated version of the Treaty on the Functioning of the European Union (TFEU) (2016) *Official Journal of the European Union C* 202/1. Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12016E/TXT** (Accessed 29 April 2022).

Cooper, K., Quayle, E., Jonsson, L. and Svedin, C. G. (2016) 'Adolescents and self-taken sexual images: A review of the literature', *Computers in Human Behavior*, 55, part B, pp.706-716. **https://doi.org/10.1016/j.chb.2015.10.003**.

Council of Europe Lanzarote Committee (2022) *Implementation report – The Protection of Children Against Sexual Exploitation and Sexual Abuse Facilitated by Information and Communication Technologies (ICTs) – Addressing the challenges raised by child self-generated sexual images and/or videos*. Available at: **https://rm.coe.int/implementation-report-on-the-2nd-monitoring-round-the-protection-of-ch/1680a619c4** (Accessed 11 May).

Crenshaw, K. (1989) 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', *University of Chicago Legal Forum*, 1989(1). Available at: **https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf** (Accessed 7 May 2022).

Cyber Safe Kids (2021) *Annual Report 2020.* Available at: **https://www.cybersafekids.ie/wp-content/uploads/2021/09/CSK_Annual_Report_2020_web.pdf** (Accessed 18 April 2022).

C3ai (n.d.) *Glossary*. Available at: **https://c3.ai/glossary/data-science/classifier/** (Accessed 25 April 2022).

Davis, K. and Koepke, L. (2015) 'Risk and protective factors associated with cyberbullying: Are relationships or rules more protective?'. *Learning, Media and Technology*, 41(4), pp.521-545. **https://doi.org/10.1080/17439884.2014.994219**.

De Santisteban, P. and Gámez-Guadix, M. (2017) 'Prevalence and Risk Factors Among Minors for Online Sexual Solicitations and Interactions With Adults'. *The Journal of Sex Research*, 55(7), pp.939-950. **https://doi.org/10.1080/00224499.2017.1386763**.

Díaz-Aguado, M. J. and Martínez-Arias, R. (2022) 'Types of Male Adolescent Violence Against Women in Three Contexts: Dating Violence Offline, Dating Violence Online, and Sexual Harassment Online Outside a Relationship'. *Frontiers in Psychology*, 13, 850897. **https://doi.org/10.3389/fpsyg.2022.850897**.

Directive (EU) 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce') (2000) *Official Journal of the European Union L* 178. Available at: **https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32000L0031** (Accessed 21 April 2022).

Directive (EU) 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety (2001) *Official Journal L* 011. Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02001L0095-20100101** (Accessed 22 April 2022).

Directive (EU) 2009/48/EC of the European Parliament and of the Council of 18 June 2009 on the safety of toys (2009) *Official Journal of the European Union L* 170. Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02009L0048-20210521** (Accessed 22 April 2022).

Directive (EU) 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) (2010) *Official Journal L* 095. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02010L0013-20181218 (Accessed 22 April 2022).

Directive (EU) 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA (2011) *Official Journal of the European Union L* 335. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02011L0093-20111217 (Accessed 21 April 2022).

Donaldson, S., Davidson, J. and Aiken, M. (2021) *Safer technology, safer users: The UK as a world-leader in Safety Tech*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/974414/Safer_technology__safer_users-_The_UK_as_a_world-leader_in_Safety_Tech_V2.pdf (Accessed 27 April 2022).

Economist Impact and WeProtect Global Alliance (2022) *Estimates of childhood exposure to online sexual harms and their risk factors*. Available at: https://www.weprotect.org/economist-impact-global-survey/#report (Accessed 18 April 2022).

ECPAT International (2016) *Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse (commonly known as the 'Luxembourg Guidelines')*. Available at: https://ecpat.org/wp-content/uploads/2021/05/Terminology-guidelines-396922-EN-1.pdf (Accessed 18 April 2022).

ECPAT International (2020) *Summary Paper on Online Sexual Exploitation of Children*. Available at: https://ecpat.org/wp-content/uploads/2021/05/ECPAT-Summary-paper-on-Online-Child-Sexual-Exploitation-2020.pdf (Accessed 21 April 2022).

ECPAT International (2021) *Global Boys Initiative: A global review of existing literature on the sexual exploitation of boys*. Available at : https://ecpat.org/wp-content/uploads/2021/09/Global-Boys-Initiative-Literature-Review-ECPAT-International-2021.pdf (Accessed 7 May 2022).

ECPAT International and WeProtect Global Alliance (2022) *Child Sexual Abuse and Exploitation Online: Survivors Perspectives*. Available at: https://www.weprotect.org/survivors-perspectives/ (Accessed 18 April 2022).

Ekambaranathan, A., Zhao, J. and Van Kleek, M. (2021) '"Money makes the world go around": Identifying Barriers to Better Privacy in Children's Apps From Developers' Perspectives'. *CHI Conference on Human Factors in Computing Systems* (CHI '21), Yokohama, Japan. 8-13 May. ACM, New York, United States, Article 46, pp.1-15. https://doi.org/10.1145/3411764.3445599.

Endendijk, J. J., Deković, M., Vossen, H., van Baar, A. L. and Reitz, E. (2022) 'Sexual Double Standards: Contributions of Sexual Socialization by Parents, Peers, and the Media'. *Archives of Sexual Behavior*, 51, pp.1721-1740. https://doi.org/10.1007/s10508-021-02088-4.

End Violence Against Children (n.d.), *Safe Online*. Available at: https://www.end-violence.org/safe-online (Accessed: 18 April 2022).

Equality Now, TrustLaw and Thomson Reuters Foundation (2021) *Ending Online Sexual Exploitation and Abuse of Women and Girls: A Call for International Standards*. Available at: https://equalitynow.storage.googleapis.com/wp-content/uploads/2021/11/13160619/Ending-OSEA-Report.pdf (Accessed 25 April 2022).

Erickson, L. B., Wisniewski, P., Xu, H., Carroll, J. M., Rosson, M. B. and Perkins, D. F. (2016) 'The boundaries between: Parental involvement in a teen's online world', *Journal of the Association for Information Science and Technology*, 67(6), pp.1384-1403. https://doi.org/10.1002/asi.23450.

eSafety Commissioner (2019) *Principles and background*. Available at: https://www.esafety.gov.au/industry/safety-by-design/principles-and-background (Accessed 18 April 2022).

eSafety Commissioner (n.d.) *Young people and social media usage*. Available at: https://www.esafety.gov.au/research/youth-digital-dangers/social-media-usage (Accessed 18 April 2022).

European Commission (2020) *EU strategy for a more effective fight against child sexual abuse*. Available at: https://ec.europa.eu/home-affairs/system/files/2020-07/20200724_com-2020-607-commission-communication_en.pdf (Accessed 22 April 2022).

European Commission (2021) *Proposal for a Regulation of The European Parliament and of The Council on general product safety, amending Regulation (EU) No 1025/2012*. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0346&qid=1628522210573 (Accessed 22 April 2022).

European Commission (n.d.) *Areas of EU action*. Available at: https://ec.europa.eu/info/about-european-commission/what-european-commission-does/law/areas-eu-action_en#EU (Accessed 18 April 2022).

European Commission (n.d.) C*ombating child sexual abuse – review of EU rules*. Available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13073-Combating-child-sexual-abuse-review-of-EU-rules_en (Accessed 21 April 2022).

European Commission (n.d.) *Combating gender-based violence – protecting victims and punishing offenders*. Available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12682-Combating-gender-based-violence-protecting-victims-and-punishing-offenders_en
(Accessed 21 April 2022).

European Commission (n.d.) *EU action on the rights of the child*. Available at: https://ec.europa.eu/info/policies/justice-and-fundamental-rights/rights-child/eu-action-rights-child_en (Accessed 18 April 2022).

European Commission (n.d.) *Fighting child sexual abuse: detection, removal and reporting of illegal content online*. Available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12726-Fighting-child-sexual-abuse-detection-removal-and-reporting-of-illegal-content-online_en (Accessed 21 April 2022).

European Commission (n.d.) *Legal framework to protect children*. Available at: https://ec.europa.eu/home-affairs/policies/internal-security/child-sexual-abuse/legal-framework-protect-children_en
(Accessed 22 April 2022).

European Parliament (2012) *EU Framework of Law for Children's Rights*. Available at: https://www.europarl.europa.eu/RegData/etudes/note/join/2012/462445/IPOL-LIBE_NT(2012)462445_EN.pdf
(Accessed 18 April 2022).

European Parliament (n.d.) *Legislative Train Schedule – A Europe Fit for the Digital Age*. Available at: https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-digital-services-act
(Accessed 21 April 2022).

European Union Agency for Fundamental Rights (n.d.) *Child rights impact assessment*. Available at: https://fra.europa.eu/en/content/child-rights-impact-assessment#:~:text=Child%20rights%20impact%20assessment%20is,development%20of%20policies%20and%20laws (Accessed 25 April 2022).

Fiesler, C., Dye, M., Feuston, J. L, Hiruncharoenvate, C., Hutto,C. J., Morrison, S., Khanipour Roshan, P., Pavalanathan, U., Bruckman, A. S., De Choudhury, M. and Gilbert, E. (2017) 'What (or Who) Is Public? Privacy Settings and Social Media Content Sharing', *ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*, Portland, United States, 25 February – 1 March. ACM, New York, United States, pp.567–580. https://doi.org/10.1145/2998181.2998223.

Finkelhor, D., Jones, L. and Mitchell, K. (2021) ‘Teaching Privacy: A flawed strategy for children’s online safety’. *Child Abuse & Neglect*, 117. https://doi.org/10.1016/j.chiabu.2021.105064.

5Rights Foundation (2018) *Disrupted Childhood*. Available at: https://5rightsfoundation.com/static/5Rights-Disrupted-Childhood.pdf (Accessed 25 April 2022).

5Rights Foundation (2021) *Pathways: how digital design puts children at risk*. Available at: https://5rightsfoundation.com/uploads/Pathways-how-digital-design-puts-children-at-risk.pdf (Accessed 18 April 2022).

5Rights Foundation (2021) *Toy Safety Directive Review*. Available at: https://5rightsfoundation.com/uploads/ToySafetyDirectiveReview-5RightsRecommendations.pdf (Accessed 25 April 2022).

5Rights Foundation (n.d.) *In our own words – children’s rights in the digital world*. Available at https://5rightsfoundation.com/In_Our_Own_Words_Young_Peoples_Version_Online.pdf (Accessed 22 April 2022).

5Rights Foundation (n.d.) *Risky-by-Design*. Available at: https://www.riskyby.design/friend-suggestions (Accessed 18 April 2022).

Gallagher, K. E. and Parrott, D. J.(2011) ‘What accounts for men’s hostile attitudes toward women? The influence of hegemonic male role norms and masculine gender role stress’, *Violence Against Women*, 17(5), pp.568-583. https://doi.org/10.1177/1077801211407296.

Ghosh, A. K., Badillo-Urquiola, K., Guha, S., LaViola Jr., J. J. and Wisniewski, P. (2018) ‘Safety vs. Surveillance: What Children Have to Say about Mobile Apps for Parental Control’. *CHI Conference on Human Factors in Computing Systems (CHI ’18)*, Montréal, Canada, 21-26 April, pp.1-14. https://doi.org/10.1145/3173574.3173698.

Ghosh, A. K., Hughes, C. E. and Wisniewski, P. (2020) ‘Circle of Trust: A New Approach to Mobile Online Safety for Families’, *2020 CHI Conference on Human Factors in Computing Systems*. Honolulu, United States, 25-30 April. ACM, New York, United States, pp.1-14. https://doi.org/10.1145/3313831.3376747.

Gillespie, S. M., Bailey, A., Squire, T., Carey, M. L., Eldridge, H. J. and Beech, A. R. (2018) ‘An Evaluation of A Community-Based Psycho-Educational Program for Users of Child Sexual Exploitation Material’, *Sexual Abuse*, 30(2), pp.169-191. https://doi.org/10.1177/1079063216639591.

Gray, C. M., Santos, C, Bielova, N., Toth, M. and Clifford, D. (2021) ‘Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective’, *CHI Conference on Human Factors in Computing Systems (CHI ’21)*, Yokohama, Japan, 8-13 May. ACM, New York, United States. https://doi.org/10.1145/3411764.3445779.

Habib, H., Pearman, S., Wang, J., Zou, Y., Acquisti, A., Cranor, L., Sadeh, N. and Schaub, F. (2020) ‘“It’s a scavenger hunt”: Usability of Websites’ Opt-Out and Data Deletion Choices’. *CHI Conference on Human Factors in Computing Systems (CHI ’20)*, Honolulu, United States, 25-30 April. ACM, New York, United States. https://doi.org/10.1145/3313831.3376511.

Hao, K. (2021) ‘The Facebook whistleblower says its algorithms are dangerous. Here’s why.’, *MIT Technology Review*, 5 October. Available at: https://www.technologyreview.com/2021/10/05/1036519/facebook-whistleblower-frances-haugen-algorithms/ (Accessed 27 April 2022).

Hartikainen, H., Razi, A. and Wisniewski, P. (2021) ‘Safe Sexting: The Advice and Support Adolescents Receive from Peers Regarding Online Sexual Risks’, *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 42, pp.1-31 https://doi.org/10.1145/3449116.

Hasebrink, U., Livingstone, S. and Haddon, L. (2008) *Comparing children's online opportunities and risks across Europe: cross-national comparisons for EU Kids Online*. Available at: http://eprints.lse.ac.uk/21656/1/D3.2_Report-Cross_national_comparisons.pdf (Accessed 27 April 2022).

INHOPE (2021) *Annual Report 2020*. Available at: **https://inhope.org/media/pages/the-facts/download-our-whitepapers/annual-report/bb4dd3cdc3-1628156678/inhope-annual-report-2020.pdf** (Accessed 18 April 2022).

INHOPE (2021) *What is self-generated CSAM?* Available at: **https://www.inhope.org/EN/articles/what-is-self-generated-csam** (Accessed 18 April 2022).

INHOPE (2021) *What is sexting?* Available at: **https://inhope.org/EN/articles/what-is-sexting?** (Accessed 27 April 2022).

Information Commissioner's Office (2021) *Age Assurance for the Children's Code*. Available at: **https://ico.org.uk/media/about-the-ico/documents/4018659/age-assurance-opinion-202110.pdf** (Accessed 25 April 2022).

Insoll, T., Ovaska, A. and Vaaranen-Valkonen, N. (2021) *CSAM Users in the Dark Web: Protecting Children Through Prevention*. Available at: **https://drive.google.com/file/d/1EUBsU0A8XYw8QNUg3JKqemIRoLO9cYPt/view** (Accessed 18 April 2022).

Instagram (2021) *Giving Young People a Safer, More Private Experience*, 27 June. Available at: **https://about.instagram.com/blog/announcements/giving-young-people-a-safer-more-private-experience** (Accessed 7 May 2022).

Instagram (n.d.) *Privacy settings and information*. Available at: **https://help.instagram.com/196883487377501**. (Accessed 7 May 2002).

Internet Watch Foundation (2020) *Face the facts – The Annual Report 2020*. Available at: **https://annualreport2020.iwf.org.uk/** (Accessed 18 April 2022).

Internet Watch Foundation (2020) *Trend: 'Self-generated' content*. Available at: **https://annualreport2020.iwf.org.uk/trends/international/selfgenerated** (Accessed 25 April 2022).

Internet Watch Foundation (2021) *'Appalling' rise of 'devious' criminals tricking children into sexually abusing themselves on camera*. Available at: **https://www.iwf.org.uk/news-media/news/appalling-rise-of-devious-criminals-tricking-children-into-sexually-abusing-themselves-on-camera/** (Accessed 18 April 2022).

Internet Watch Foundation (2022) *Three-fold increase of abuse imagery of 7-10-year-olds as IWF detects more child sexual abuse material online than ever before*. 13 January. Available at: **https://www.iwf.org.uk/news-media/news/three-fold-increase-of-abuse-imagery-of-7-10-year-olds-as-iwf-detects-more-child-sexual-abuse-material-online-than-ever-before/** (Accessed 18 April 2022).

Internet Watch Foundation (2022) *The Annual Report 2021*. Available at: **https://annualreport2021.iwf.org.uk/trends/** (Accessed 6 May 2022).

Jia, H., Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Risk-taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors', *ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '15)*, Vancouver, Canada, 14-18 March. ACM, New York, United States, pp.583-599. **https://doi.org/10.1145/2675133.2675287**.

Joeckel, S., Dogruel, L. (2019) 'Default effects in app selection: German adolescents' tendency to adhere to privacy or social relatedness features in smartphone apps', *Mobile Media & Communication*, 8(1), pp.22-41. **https://doi.org/10.1177/2050157918819616** (Accessed 8 May 2022).

Joleby, M., Lunde, C., Landström, S. and Jonsson, L. S. (2020) '"All of Me Is Completely Different": Experiences and Consequences Among Victims of Technology-Assisted Child Sexual Abuse', *Frontiers in Psychology*, 11. https://doi.org/10.3389/fpsyg.2020.606218.

Jonsson, L. S., Fredlund, C., Priebe, G., Wadsby, M. and Svedin, C. G. (2019) 'Online sexual abuse of adolescents by a perpetrator met online: a cross-sectional study', *Child and Adolescent Psychiatry and Mental Health*, 13, 32. https://doi.org/10.1186/s13034-019-0292-1.

Kaushal, R., Saha, S., Bajaj, P. and Kumaraguru, P. (2016) *KidsTube: Detection, characterization and analysis of child unsafe content & promoters on YouTube*. https://doi.org/10.48550/arXiv.1608.05966.

Keller, M. and Dance, G. (2019) 'Images of Child Sexual Abuse. What Went Wrong?', *New York Times*, 28 September. Available at: https://www.nytimes.com/interactive/2019/09/28/us/child-sex-abuse.html (Accessed 18 April 2022).

Laghi, F. and Schneider, B. (2013) 'Knowing when not to use the Internet: Shyness and adolescents' on-line and off-line interactions with friends'. *Computers in Human Behavior*, 29, pp.51-57. https://doi.org/10.1016/j.chb.2012.07.015.

Laird, J., Klettke, B., Hall, K., Clancy, E. and Hallford, D. (2020) *'Demographic and Psychosocial Factors Associated With Child Sexual Exploitation – A Systematic Review and Meta-analysis'*, Jama Network Open, 3(9). https://doi.org/10.1001/jamanetworkopen.2020.17682.

'La Quadrature du Net and others v Premier Ministre and others' (2020) C-511/18, C-512/18 and C-520/18, ECLI:EU:C:2020:791. Available at: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:62018CJ0511 (Accessed 22 April 2022).

Letourneau, E. J., Schaeffer, C. M., Bradshaw, C. P. and Feder, K. A. (2017) 'Preventing the Onset of Child Sexual Abuse by Targeting Young Adolescents With Universal Prevention Programming', *Child Maltreatment*, 22(2), pp.100-111. https://doi.org/10.1177/1077559517692439.

Livingstone, S., Blum-Ross, A., Pavlick, J. and Ólafsson, K. (2018) *In the digital home, how do parents support their children and who supports them?* Available at: https://www.lse.ac.uk/media-and-communications/assets/documents/research/preparing-for-a-digital-future/P4DF-Survey-Report-1-In-the-digital-home.pdf (Accessed 25 April 2022).

Livingstone, S. and Helsper, E. (2010) 'Balancing opportunities and risks in teenagers' use of the internet: the role of online skills and internet self-efficacy'. *New Media & Society*. 12(2), pp.309-329. https://doi.org/10.1177/1461444809342697.

Livingstone, S., Mascheroni, G., Dreier, M., Chaudron, S. and Lagae, K. (2015) *How parents of young children manage digital devices at home: The role of income, education and parental style*. Available at: http://eprints.lse.ac.uk/63378/1/__lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_EU%20Kids%20Online_EU_Kids_Online_How%20parents%20manage%20digital%20devices_2016.pdf (Accessed 25 April 2022).

Livingstone, S. and Mason, J. (2015) *Sexual rights and sexual risks among youth online*. Available at: http://eprints.lse.ac.uk/id/eprint/64567 (Accessed 11 May).

Livingstone, S., Ólafsson, K., Helsper, E. J., Lupianez-Villanueva, F., Veltri, G. A. and Folkvord, F. (2017) 'Maximizing opportunities and minimizing risks for children online: the role of digital skills in emerging strategies of parental mediation', *Journal of Communication*, 67(1), pp.82-105. https://doi.org/10.1111/jcom.12277.

Livingstone, S. and Smith, P. K. (2014) 'Annual Research Review: Harms experienced by child users of online and mobile technologies: the nature, prevalence and management of sexual and aggressive risks in the digital age', *The Journal of Child Psychology and Psychiatry*, 55(6), pp.635-54. **https://doi.org/10.1111/jcpp.12197**.

Livingstone, S., Stoilova, M. and Nandagiri, R. (2019) *Children's data and privacy online: Growing up in a digital age*. An evidence review. Available at: **https://www.lse.ac.uk/media-and-communications/assets/documents/research/projects/childrens-privacy-online/Evidence-review.pdf** (Accessed 29 April 2022).

Madigan, S., Villani, V., Azzopardi, C., Laut, D., Smith, T., Temple, J. R., Browne, D., Dimitropoulos, G. (2018) 'The Prevalence of Unwanted Online Sexual Exposure and Solicitation Among Youth: A Meta-Analysis', *The Journal of Adolescent Health*, 63 (2), pp 133-141. **https://doi.org/10.1016/j.jadohealth.2018.03.012**.

Marley, R. (2021) 'Age Verification for Social Media- Protecting the Younger Victims of Online Scams', *ShuftiPro,* 14 October. Available at: **https://shuftipro.com/blog/age-verification-for-social-media-protecting-the-younger-victims-of-online-scams/** (Accessed 25 April 2022).

McHugh, B. C., Wisniewski, P., Rosson, M. B. and Carroll, J. M. (2018) 'When social media traumatizes teens: The roles of online risk exposure, coping, and post-traumatic stress', *Internet Research*, 28(5), pp.1169-1188. **https://doi.org/10.1108/IntR-02-2017-0077**.

Meridan, H., Thakker, J., Wilson, N. and Boer, D. (2011) 'Assessing the internal structure of the COPINE scale', *Psychology*, Crime and Law, 19(1), pp.21-34. **https://doi.org/10.1080/1068316X.2011.598158**.

Merino, E., Díaz-Aguado, M. J., Falcón, L. and Martínez-Arias, R. (2021) 'Masculine Gender Role Stress as a Mediator of the Relationship Between Justification of Dominance and Aggression and Male Adolescent Dating Violence Against Women', *Psicothema*, 33(2), pp.206-213. **https://doi.org/10.7334/psicothema2020.275**.

Meta Transparency Center (2022) *Child Sexual Exploitation, Abuse and Nudity*. Available at: **https://transparency.fb.com/en-gb/policies/community-standards/child-sexual-exploitation-abuse-nudity/** (Accessed 25 April 2022).

Meta (2018) *Facebook: New Technology to Fight Child Exploitation*. 24 October. Available at: **https://about.fb.com/news/2018/10/fighting-child-exploitation/** (Accessed 27 April 2022).

Meta (2021) *Instagram: Asking People for Their Birthday on Instagram*. 30 August. Available at: **https://about.fb.com/news/2021/08/asking-people-for-their-birthday-on-instagram/** (Accessed 27 April 2022).

Meta (2021) *Instagram: Continuing to Make Instagram Safer for the Youngest Members of Our Community*. 16 March. Available at: **https://about.fb.com/news/2021/03/continuing-to-make-instagram-safer-for-the-youngest-members-of-our-community/** (Accessed 25 April 2022).

Meta (2022) *Facebook Community Standards – Child Sexual Exploitation, Abuse and Nudity*. Available at: **https://transparency.fb.com/en-gb/policies/community-standards/child-sexual-exploitation-abuse-nudity/** (Accessed 25 April 2022).

Mhaidli, A.H., Zou, Y. and Schaub, F. (2019) '"We Can't Live Without Them!" App Developers' Adoption of Ad Networks and Their Considerations of Consumer Risks' *Fifteenth USENIX Symposium on Usable Privacy and Security*. Santa Clara, United States, 12-13 August, pp.225-244. Available at: **https://www.usenix.org/system/files/soups2019-mhaidli.pdf** (Accessed 25 April 2022).

Milkaite, I. and Lievens, E. (2020) 'Child-friendly transparency of data processing in the EU: from legal requirements to platform policies'. *Journal of Children and Media*, 14(1), pp.5-21. **https://doi.org/10.1080/17482798.2019.1701055**.

Mishna, F., Milne, E., Cook, C., Slane, A. and Ringrose, J. (2021) 'Unsolicited Sexts and Unwanted Requests for Sexts: Reflecting on the Online Sexual Harassment of Youth', *Youth & Society*. https://doi.org/10.1177/0044118X211058226.

Muñoz, F., Isaza, G., Castillo, L. (2021) 'SMARTSEC4COP: Smart Cyber-Grooming Detection Using Natural Language Processing and Convolutional Neural Networks'. https://doi.org/10.1007/978-3-030-53036-5_2.

Namy, S., Carlson, C., O'Hara, K., Nakuti, J., Bukuluki, P., Lwanyaaga, J., Namakula, S., Nanyunja, B., Wainberg, M., Naker, D. and Michau, L. (2017) 'Towards a feminist understanding of intersecting violence against women and children in the family', *Social Science & Medicine*, Vol. 184, pp.40-48. https://doi.org/10.1016/j.socscimed.2017.04.042.

Nash, V., Adler, J. R., Horvath, M. A. H., Livingstone, S., Marston, C., Owen, G. and Wright, J. (2015) *Identifying the routes by which children view pornography online: implications for future policy-makers seeking to limit viewing*. Available at: http://eprints.lse.ac.uk/65450/1/__lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_Livingstone%2C%20S_Identifying%20the%20routes_Livingstone_Identifying%20the%20routes_2016.pdf (Accessed 22 April 2022).

Nash, V., O'Connell, R., Zevenbergen, B. and Mishkin, A. (December 2012-December 2013) *Effective Age Verification Techniques: Lessons to Be Learnt from the Online Gambling Industry*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2658038 (Accessed 25 April 2022).

NCMEC (2019) *Captured on film: survivors of child sex abuse imagery are stuck in a unique cycle of trauma*. Available at: https://calio.org/wp-content/uploads/2020/03/Captured-on-Film-Survivors-of-Child-Sex-Abuse-Imagery-are-Stuck-in-a-Unique-Cycle-of-Trauma.pdf (Accessed 18 April 2022).

NCMEC (2021) *CyberTipline 2021 Report*. Available at: https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata (Accessed 7 May 2022).

NCMEC (2022) *Sextortion: The Hidden Pandemic*. Available at: https://www.missingkids.org/blog/2022/sextortion-the-hidden-pandemic (Accessed 18 April 2022).

Nealon, L. (2021) 'TikTok Takes Fighting Sexploitation Seriously', *National Center on Sexual Exploitation*. 14 July. Available at: https://endsexualexploitation.org/articles/tiktok-takes-fighting-sexploitation-seriously/ (Accessed 27 April 2022).

NetClean (2019) *NetClean Report 2018 – A Report about child sexual abuse*. Available at: https://www.netclean.com/wp-content/uploads/2018/12/The-NetClean-Report-2018_Web.pdf (Accessed 25 April 2022).

NetClean (2019) *NetClean Report 2019 – A Report about child sexual abuse crime*. Available at: https://www.netclean.com/netclean-report-2019/ (Accessed 18 April 2022).

NetClean (2021) *NetClean Report – COVID-19 Impact 2020 – A Report about child sexual abuse crime*. Available at: https://www.netclean.com/wp-content/uploads/2021/01/NetCleanReport_COVID19_Impact2020_pages.pdf (Accessed 18 April 2022).

New York Times (2020) *A Third of TikTok's U.S. Users May Be 14 or Under, Raising Safety Questions*. 14 August. Available at: https://www.nytimes.com/2020/08/14/technology/tiktok-underage-users-ftc.html (Accessed 18 April 2022).

Nezhad, M. M. and Mehrnezhad, M. (2018) 'A child recognition system based on image selection patterns', *7th Workshop on Socio-Technical Aspects in Security and Trust (STAST '17)*, Orlando, United States, 5 December. ACM, New York, United States, pp. 76-81. https://doi.org/10.1145/3167996.3168003.

Norwegian Consumer Council (2018) *Deceived by design: How tech companies use dark patterns to discourage us from exercising our rights to privacy*. Available at: https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf (Accessed 25 April 2022).

NSPCC (2021) *New figures reveal four in five victims of online grooming crimes are girls*. Available at: https://www.nspcc.org.uk/about-us/news-opinion/2021/online-grooming-crimes-girls/ (Accessed 18 April 2022).

NSPCC (2021) *Record high number of recorded grooming crimes lead to calls for stronger online safety legislation*. 24 August. Available at: https://www.nspcc.org.uk/about-us/news-opinion/2021/online-grooming-record-high/ (Accessed 18 April 2022).

NSPCC (2021) *Statistics briefing: harmful sexual behaviour.* Available at: https://learning.nspcc.org.uk/media/1661/statistics-briefing-harmful-sexual-behaviour.pdf. (Accessed 25 April 2022).

Odink, I. (2019) *Children's Rights in the EU: Marking 30 Years of the UN Convention on the Rights of the Child*. European Parliamentary Research Service. Available at: https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/644175/EPRS_BRI(2019)644175_EN.pdf (Accessed 21 April 2022).

OECD (2021) *Children in the digital environment: Revised typology of risks – OECD Digital Economy Papers*, No. 302, Figure 1, p.7. https://doi.org/10.1787/9b8f222e-en.

Orben, A., Przybylski, A.K., Blakemore, S.J. and Kievit, R. A. (2022) 'Windows of developmental sensitivity to social media', *Nature Communications*, 13, 1649. https://doi.org/10.1038/s41467-022-29296-3.

OSCE Office of the Special Representative and Co-ordinator for Combating Trafficking in Human Beings and Tech Against Trafficking (2020) *Leveraging innovation to fight trafficking in human beings: A comprehensive analysis of technology tools*. Available at: https://www.osce.org/files/f/documents/9/6/455206_1.pdf (Accessed: 18 April 2022).

Oversight Board (2021) *Oversight Board publishes transparency report for third quarter of 2021*. Available at: https://www.oversightboard.com/news/640697330273796-oversight-board-publishes-transparency-report-for-third-quarter-of-2021/ (Accessed 7 May 2022).

Park, Y., Gentile, D. A., Morgan, J., He, L., Allen, J. J., Jung, S. M., Chua, J., Koh, A. (2020) *Child Online Safety Index Report*, p.15. Available at: https://www.dqinstitute.org/wp-content/uploads/2020/02/2020-COSI-Findings-and-Methodology-Report.pdf (Accessed: 18 April 2022).

Pasquale, L., Zippo, P., Curley, C., O'Neill, B and Mongiello, M. (2022) 'Digital Age of Consent and Age Verification: Can They Protect Children?', *IEEE Software*, 39(3), pp.50-57. https://doi.org/10.1109/MS.2020.3044872.

Patchin, J. W. and Hinduja, S. (2018) 'Sextortion Among Adolescents: Results From a National Survey of U.S. Youth', *Sexual Abuse*, 32(1), pp.30-51. https://doi.org/10.1177/1079063218800469.

Patchin, J. W. and Hinduja, S. (2020) 'It is time to teach safe texting', *Journal of Adolescent Health*, 66(2), pp.140-143. https://doi.org/10.1016/j.jadohealth.2019.10.010.

Perkins, D., Merdian, H., Schumacher, B., Bradshaw, H. and Stevanovic, J. (2018) *Interventions for perpetrators of online child sexual exploitation – A scoping review and gap analysis.* Available at: https://www.csacentre.org.uk/documents/online-cse-interventions/ (Accessed 11 May).

Phippen, A. and Brennan, M. (2021) *Child Protection and Safeguarding Technologies Appropriate or Excessive 'Solutions' to Social Problems?* New York: Routledge.

Pinter, A. T., Wisniewski, P., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Adolescent Online Safety: Moving Beyond Formative Evaluations to Designing Solutions for the Future', *Conference on Interaction Design and Children (IDC '17)*, Stanford, United States, 27-30 June. ACM, New York, United States, pp.352-357. **https://doi.org/10.1145/3078072.3079722**.

Prichard, J., Wortley, R., Watters, P.A., Spiranovic, C., Hunn, C. and Krone T. (2022) 'Effects of Automated Messages on Internet Users Attempting to Access "Barely Legal" Pornography', *Sexual Abuse*, 34(1), pp.106-124. **https://doi.org/10.1177/10790632211013809**.

Putman, L. (2022) 'Facebook Has a Child Predation Problem', *Wired*, 13 March. Available at: **https://www.wired.com/story/facebook-has-a-child-predation-problem/** (Accessed 7 May 2022).

Quayle, E. and Koukopoulus, N. (2018) 'Deterrence of Online Child Sexual Abuse and Exploitation', *Policing*, p.13(3), pp.345-362. **https://doi.org/10.1093/police/pay028**.

Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2020) 'Let's Talk about Sext: How Adolescents Seek Support and Advice about Their Online Sexual Experiences', *2020 CHI Conference on Human Factors in Computing Systems*, Honolulu, United States, 25-30 April. ACM, New York, United States. **https://doi.org/10.1145/3313831.3376400**.

Razi, A., Kim, S., Alsoubai, A., Stringhini, H., Solorio, T., De Choudhury, M. and Wisniewski, P. (2021) 'A Human-Centered Systematic Literature Review of the Computational Approaches for Online Sexual Risk Detection', *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 465, pp.1-38. **https://doi.org/10.1145/3479609**.

Regulation (EU) 2016/679 of The European Parliament and of The Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (2016) *Official Journal of the European Union L* 119/1. Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=DA** (Accessed 22 April 2022).

R v. Oliver (2002) EWCA Crim, case 2766. Available at: **https://vlex.co.uk/vid/r-v-oliver-r-792617673** (Accessed 25 April 2022).

Salter, M. and Wong, T. (2021) *Research report – The impact of COVID-19 on the risk of online child sexual exploitation and the implications for child protection*, p.38. Available at: **https://www.arts.unsw.edu.au/sites/default/files/documents/eSafety-OCSE-pandemic-report-salter-and-wong.pdf** (Accessed 18 April 2022).

Satariano, A. (2021) 'Facebook's oversight board faults its policy on preferential treatment', *New York Times*, 21 October. Available at: **https://www.nytimes.com/2021/10/21/business/facebook-oversight-board-members-criticism.html** (Accessed 7 May 2022).

Schoeps, K., Peris Hernández, M. P., Garaigordobil, M. and Montoya-Castilla, I. (2020) 'Risk factors for being a victim of online grooming in adolescents'. *Psicothema*, 32(1), pp.15-23. **https://doi.org/10.7334/psicothema2019.179**.

Sharma, M.K., Anand, N., Kumar, K., Lenin Singh, R., Thakur, P.C., Mondal, I., Kohli, T. (2021) 'Constructing the understanding of teenagers deviant use of cyberspace'. *International Journal of Social Psychiatry*. 67(8), pp.1068-1071. **https://doi.org/10.1177/0020764020975791**.

Smahel, D., Machackova, H., Mascheroni, G., Dedkova, L., Staksrud, E., Ólafsson, K., Livingstone, S. and Hasebrink, U. (2020) *EU Kids Online 2020: Survey results from 19 countries.* **https://doi.org/10.21953/lse.47fdeqj01ofo**.

Smirnova, S., Livingstone, S. and Stoilova, M. (2021) *Understanding of user needs and problems: a rapid evidence review of age assurance and parental controls*. Available at: http://eprints.lse.ac.uk/112559/ (Accessed 22 April 2022).

Smith, P. K. and Livingstone, S. (2017) 'Child Users of Online and Mobile Technologies – Risks, Harms and Intervention', *Child Psychology and Psychiatry: Frameworks for Clinical Training and Practice*, pp.141-148. https://doi.org/10.1002/9781119170235.ch17.

Smith, P. K., Thompson, F. and Davidson, J. (2014) 'Cyber safety for adolescent girls: bullying, harassment, sexting, pornography, and solicitation', *Current Opinion in Obstetrics and Gynecology*, 26(5), pp.360-365. https://doi.org/10.1097/GCO.0000000000000106.

Snap Inc. (2022) *Community Guidelines*. Available at: https://snap.com/en-US/community-guidelines (Accessed 25 April 2022).

Snap Inc. (n.d.) *Privacy settings*. Available at: https://support.snapchat.com/en-GB/a/privacy-settings2 (Accessed 7 May 2022).

Snap Inc. (n.d.) *Snapchat Safety Center*. Available at: https://snap.com/en-US/safety/safety-center (Accessed 27 April 2022).

Sneddon, H., Gojkovic Grimshaw, D., Livingstone, N. and Macdonald, G. (2020) 'Cognitive-behavioural therapy (CBT) interventions for young people aged 10 to 18 with harmful sexual behaviour', *Cochrane Database of Systematic Reviews*, 6, CD009829. https://doi.org/10.1002/14651858.CD009829.pub2.

SRHR Africa Trust, TrustLaw, Arnold & Porter Kaye Scholer LLP (2018) *Age of Consent: Global Legal Review*, p.12. Available at: https://www.satregional.org/wp-content/uploads/2018/05/Age-of-consent-Global-Legal-Review.pdf (Accessed 11 May).

Stoilova, M., Livingstone, S. and Khazbak, R. (2021) *Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children's internet use and outcomes*. Available at: https://www.unicef-irc.org/publications/pdf/Investigating-Risks-and-Opportunities-for-Children-in-a-Digital-World.pdf (Accessed 27 April 2022).

Sultana, S., Deb, M., Bhattacharjee, A., Hasan, S., Raihanul Alam, S. M., Chakraborty, T., Roy, P., Fairuz Ahmed, S., Moitra, A., Ashraful Amin, M., Najmul Islam, A. K. M. and Ishtiaque Ahmed, S. (2021) ''Unmochon': A Tool to Combat Online Sexual Harassment over Facebook Messenger'. *CHI Conference on Human Factors in Computing Systems (CHI '21)*, Yokohama, Japan, 8-13 May, pp.1-18. https://doi.org/10.1145/3411764.3445154.

Tahaei, M., Frik, A. and Vaniea, K. (2021) 'Deciding on Personalized Ads: Nudging Developers About User Privacy'. *USENIX Symposium on Usable Privacy and Security (SOUPS 2021)*, Virtual Event, 8-10 August. https://doi.org/10.7488/ds/3045.

Tariq, M. U., Ghosh, A. K., Badillo-Urquiola, K., Jha, A., Koppal, S. and Wisniewski, P. (2018) 'Designing light filters to detect skin using a low-powered sensor', *SoutheastCon 2018*, St. Petersburg, United States, 19-22 April. IEEE. https://doi.org/10.1109/SECON.2018.8479027.

Tariq, M. U., Razi, A., Badillo-Urquiola, K. and Wisniewski, P. (2019) 'A Review of the Gaps and Opportunities of Nudity and Skin Detection Algorithmic Research for the Purpose of Combatting Adolescent Sexting Behaviors'. *21st International Conference on Human-Computer Interaction*. Orlando, United States, 26-31 July. Springer, Cham. https://doi.org/10.1007/978-3-030-22636-7_6.

Tech Coalition (2020) *The Technology Coalition Announces Project Protect – A Plan to Combat Online Child Sexual Abuse* [10 June] Available at: https://www.technologycoalition.org/newsroom/the-tech-coalition-announces-project-protect (Accessed 22 February 2022).

The Verge (2021). Child Safety Platforms. Thorn Report. The Verge (2021) *The child safety problem on platforms is worse than we knew*, 12 May. Available at: https://www.theverge.com/2021/5/12/22432863 (Accessed 8 May 2022).

Thorn (2020) T*he road to Safer: Equipping industry to end CSAM*. Available at: https://www.thorn.org/blog/announcing-safer-built-by-thorn-eliminate-csam/ (Accessed 21 April 2022).

Thorn (2021) *Responding to Online Threats: Minors' Perspectives on Disclosing, Reporting, and Blocking*. Available at: https://info.thorn.org/hubfs/Research/Responding to Online Threats_2021-Full-Report.pdf (Accessed: 18 April 2022).

TikTok (2021) *Legal – Privacy Policy*. Available at: https://www.tiktok.com/legal/privacy-policy-eea?lang=en (Accessed 25 April 2022).

TikTok (2021) *Strengthening privacy and safety for youth on TikTok,* 13 January. Available at: https://newsroom.tiktok.com/en-us/strengthening-privacy-and-safety-for-youth (Accessed 7 May 2022).

TikTok (n.d.) *Community Guidelines – Minor safety*. Available at: https://www.tiktok.com/community-guidelines?lang=en#31 (Accessed 25 April 2022).

TikTok (n.d.) *New User Guide*. Available at: https://www.tiktok.com/safety/en/new-user-guide/ (Accessed 25 April 2022).

TikTok (n.d.) *Teen privacy and safety settings*. Available at: https://support.tiktok.com/en/account-and-privacy/account-privacy-settings/privacy-and-safety-settings-for-users-under-age-18 (Accessed 25 April 2022).

UN CRC (2021) *General comment No. 25 (2021) on children's rights in relation to the digital environment*. Available at: https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation (Accessed: 27 April 2022).

UNHCR (2018) *Report of the Special Rapporteur on the Right to Privacy*. Available at: https://digitallibrary.un.org/record/1656178/files/A_HRC_37_62-EN.pdf (Accessed 27 April 2022).

UNICEF (2017) *The State of the World's Children 2017 – Children in a Digital World*. Available at: https://www.unicef.org/media/48601/file (Accessed 25 April 2022).

UNICEF (2019) *Global Kids Online Comparative report*. Available at: https://www.unicef-irc.org/publications/pdf/GKO%20Main%20Report.pdf (Accessed: 18 April 2022).

UNICEF (2019) *Growing Up in a Connected World: Understanding Children's Risks and Opportunities in a Digital Age*. Available at: https://www.unicef-irc.org/growing-up-connected (Accessed: 18 April 2022).

UNICEF (2020) *COVID-19 and its implications for protecting children online*. Available at: https://www.unicef.org/sites/default/files/2020-04/COVID-19-and-Its-Implications-for-Protecting-Children-Online.pdf (Accessed 18 April 2022).

UNICEF (2020) *Gender Dimensions of Violence Against Children and Adolescents*. Available at: https://www.unicef.org/media/92376/file/Child-Protection-Gender-Dimensions-of-VACAG-2021.pdf (Accessed 18 April 2022).

UNICEF (2021) *Digital Age Assurance Tools and Children's Rights Online across the Globe*. Available at: http://c-fam.org/wp-content/uploads/Digital-Age-Assurance-Tools-and-Children-s-Rights-Online-across-the-Globe-1_LT.pdf (Accessed 18 April 2022).

UNICEF (2021) *Ending online child sexual exploitation and abuse: Lessons learned and promising practices in low- and middle-income countries.* Available at: **https://www.unicef.org/media/113731/file/Ending%20Online%20Sexual%20Exploitation%20and%20Abuse.pdf** (Accessed 18 April 2022).

UNICEF (2021) *Policy guidance on AI for children.* Available at: **https://www.unicef.org/globalinsight/media/2356/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf** (Accessed 25 April 2022).

UNICEF Innovation, Human Rights Center, UC Berkley (2019) *Executive Summary: Artificial Intelligence and Children's Rights.* Available at: **https://www.unicef.org/innovation/media/10726/file/Executive%20Summary:%20Memorandum%20on%20Artificial%20Intelligence%20and%20Child%20Rights.pdf** (Accessed 27 April 2022).

UNICEF Innovation, Human Rights Center, UC Berkley (2019) *Memorandum on Artificial Intelligence and Child Rights.* Available at: **https://www.unicef.org/innovation/media/10501/file/Memorandum%20on%20Artificial%20Intelligence%20and%20Child%20Rights.pdf** (Accessed 27 April 2022).

United Nations (1990) *United Nations Convention on the Rights of the Child*, Article 13. Available at: **https://www.ohchr.org/sites/default/files/crc.pdf** (Accessed 27 April 2022).

van der Hof, S. (2021) 'Age assurance and age appropriate design: what is required?' *LSE Parenting for a Digital Future Blog*, 17 November. Available at: **https://blogs.lse.ac.uk/parenting4digitalfuture/2021/11/17/age-assurance/** (Accessed 25 April 2022).

van der Hof, S. and Ouburg, S. (2021) *Methods for Obtaining Parental Consent and Maintaining Children Rights.* Available at: **https://euconsent.eu/download/methods-for-obtaining-parental-consent-and-maintaining-children-rights/** (Accessed 25 April 2022).

Verification of Children Online (VoCO) project (2020) *Phase 2 Report.* Available at: **https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/934131/November_VoCO_report_V4__pdf.pdf** (Accessed 25 April 2022).

Waldman, A. E. (2020) 'Cognitive biases, dark patterns, and the 'privacy paradox'', *Current Opinion in Psychology*, 31, pp.105-109. **https://doi.org/10.1016/j.copsyc.2019.08.025**.

Walrave, M., Heirman, W. and Hallam, L. (2013) 'Under pressure to sext? Applying the theory of planned behaviour to adolescent sexting'. *Behaviour & Information Technology*, 33(1), pp.86-98. **https://doi.org/10.1080/0144929X.2013.837099**.

Wang, G., Zhao, J., Van Kleek, M. and Shadbol, N. (2021) 'Protection or punishment? relating the design space of parental control apps and perceptions about them to support parenting for online safety'. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 343, pp.1-26 **https://doi.org/10.1145/3476084**.

Weir, C., Hermann, B. and Fahl, S. (2020) From Needs to Actions to Secure Apps? *The Effect of Requirements and Developer Practices on App Security*. Available at: **https://eprints.lancs.ac.uk/id/eprint/142148** (Accessed 25 April 2022).

WeProtect Global Alliance (2021) *Global Threat Assessment 2021 – Working together to end the sexual abuse of children online.* Available at: **https://www.weprotect.org/global-threat-assessment-21/#report** (Accessed 18 April 2022).

WeProtect Global Alliance (2021) *The sexual exploitation and abuse of deaf and disabled children online.* Available at: **https://www.weprotect.org/wp-content/uploads/Intelligence-briefing-2021-The-sexual-exploitation-and-abuse-of-disabled-children.pdf** (Accessed 18 April 2022).

WeProtect Global Alliance and the Technology Coalition (2021) *Findings from WeProtect Global Alliance/ Technology Coalition survey of technology companies – Summary of findings*. Available at: **https://www.weprotect.org/survey-of-tech-companies/** (Accessed 25 April 2022).

Whittle, H., Hamilton-Giachritsis, C., Beech, A. and Collings, G. (2013) 'A Review of young people's vulnerabilities to online grooming', *Aggression and Violent Behavior*, 18, pp.135-146. **https://doi.org/10.1016/j.avb.2012.11.008**.

Wisniewski, P., Ghosh, A. K., Xu, H., Rosson, M. B. and Caroll, J. M. (2017) 'Parental Control vs. Teen Self-Regulation: Is there a middle ground for mobile online safety?', *2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. Portland, United States. 25 February – 1 March. ACM, New York, United States. **https://doi.org/10.1145/2998181.2998352**.

Wisniewski, P., Jia, H., Wang, N., Zheng, S., Xu, H., Rosson, M. B. and Carroll, J. M. (2015) 'Resilience mitigates the negative effects of adolescent internet addiction and online risk exposure'. *ACM Conference on Human Factors in Computing Systems (CHI '15)*, Seoul, South Korea, 18-23 April, pp.4029-4038. **https://doi.org/10.1145/2702123.2702240**.

Wisniewski, P., Xu, H., Rosson, M. B. and Carroll, J. M. (2017) 'Parents just don't understand: Why teens don't talk to parents about their online risk experiences', *ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*, Portland, United States, 25 February – 1 March, pp.523–540. **https://doi.org/10.1145/2998181.2998236**.

Witting, S.K. (2020) *Child sexual abuse in the digital era : Rethinking legal frameworks and transnational law enforcement collaboration*. PhD thesis. Leiden University. Available at: **http://hdl.handle.net/1887/96242** (Accessed 11 May).

World Population Review (2022) *Age of Consent by Country 2022*. Available at: **https://worldpopulationreview.com/country-rankings/age-of-consent-by-country** (Accessed 11 May).

YouTube (n.d.) *Child safety policy*. Available at: **https://support.google.com/youtube/answer/2801999?hl=en&ref_topic=9282679** (Accessed 25 April 2022).

YouTube (n.d.) **Community Guidelines**. Available at: **https://www.youtube.com/intl/en-GB/howyoutubeworks/policies/community-guidelines/#enforcing-community-guidelines** (Accessed 25 April 2022).