Terre des Hommes

CHILD SAFETY BY DESIGN

What is child safety by design?

Child safety by design is about placing the **safety of child users at the centre** of the designs of online services, whilst placing **responsibility** for child users' safety on online service providers. It includes:

- Taking **preventative steps** to ensure that known and anticipated harms have been evaluated in the design and provision of an online service;
- **Children's empowerment and autonomy** are secured as part of the in-service experience;
- Organisations take **ownership and responsibility** for child users' safety and well-being, with clear steps taken to address any issues.

UN Committee on the Rights of the Child

In its General Comment No. 25, the Committee recognises safety by design as a **necessary approach to the full protection of children's rights**, which should be integrated in the design of digital services and products that children use. Safety and protective measures should also take into account **children's evolving capacities.**

Core principles of safety by design

1. Service provider responsibility:
Service providers (like apps, platforms, and websites) must share the responsibility for online safety with its users. They are responsible and accountable for identifying and addressing known and potential harms, like through conducting risk assessments.

2. User empowerment and autonomy:
Products and services should align with the best interests of its users and designed to protect human rights. To do so, platforms and services should consult with diverse and at-risk groups, including children, to ensure their features and functions are accessible to everyone and protective of everyone's rights.

3. Transparency and accountability:
Companies should openly share data on how they enforce their policies and the effectiveness of safety features and innovations. If safety innovations are proven to work, they should be widely shared and adopted.

eSafety Commissioner Australia, Principles and background (2019)
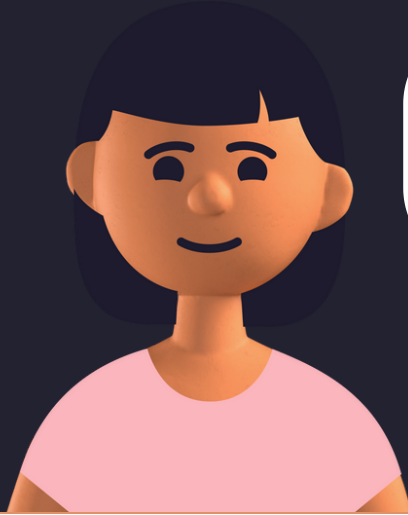
"TO MAKE CHILDREN SAFER FROM SOCIAL MEDIA, THERE SHOULD BE A SETTING TO FILTER AGE-INAPPROPRIATE CONTENTS AND THERE SHOULD BE A HELP DESK AS REPORTING MECHANISM."

Focus Group Discussions, Philippines (Child Safety by Design)

Terre des Hommes

Techniques and approaches

A child safety by design approach encompasses **specific techniques** to ensure safety (e.g. age verification, strong privacy settings by default, and hiding child 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.

Safety measures currently in place:

Various safety techniques currently exist. Though some are beneficial, they lack consistency across platforms, and some approaches are ineffective.

- **Terms and conditions** lay out the platforms' 'dos and don'ts'. These emphasise self-regulation of inappropriate behaviour, but are not enough to keep young users safe nor are they usually child-friendly.
- **Minimum age** to use platforms is set at 13 by data processing rules for most countries. Self-declaration is the usual, but ineffective, form of age verification.
- **Privacy by default** settings for children are used to protect children's content on their profiles from being publicly accessible. Functionalities used by platforms to manipulate users into sharing remains an issue.
- **Classifiers and AI technology** are used to search for and remove sexually explicit content of children or to detect predatory behaviour at high level of accuracy.
- **Intervention tactics** are also used, like warning messages.

Design solutions that can further enhance children's online safety

As well as enhancing existing safety techniques, for example through strengthening age assurance mechanisms and ensuring intelligent and proactive privacy nudges and management, solutions can also consider:

Promoting children's empowerment:
Deploying safety features that build on children's needs to help them learn the **vital skills** for safe online engagement, identify and deal with risks, and that empower them to **self-regulate** their own online behaviour, can help foster autonomy, privacy, and agency. This is especially important for adolescents during their stage of development. Providing children a safe place to get peer and/or professional support away from general (public) forums can also be an effective way to help teenagers learn and empower themselves.

Technologies that can identify risks combined with mitigation strategies:
Technologies can also help to proactively detect potential risks prior to, or in the moment they occur. Combined with risk mitigation strategies like educational and awareness prompts and nudging, these can help children be safer online. For instance, a child could get a **warning and guidance** on how best to proceed when contacted by a stranger. These prompts should be balanced, avoiding excessive strictness that urge children to avoid all online risks. Overly strict advice might cause children (especially teens) to lose interest and tune prompts out. Instead, they should guide children to appropriate action and provide tips on how to cope with these risks. Additionally, prompts can also be used to stop risky behaviour, for example, through warning both offenders and children that they might be about to post inappropriate content.

Establishing minimum legislative standards to keep children safe online:
In addition to requiring platforms to have **effective age verification methods**, **accessible reporting** and **referral mechanisms** for children, policymakers should establish minimum child online safety standards to make platforms accountable. With regulated and enforceable minimum standards and a uniform code of practice covering all elements, features and functions of a service overseen by a designated regulator, policymakers can help ensure consistency across different platforms, and hold platforms **accountable** to implement them.

For more information consult: TdH NL Child Safety by Design 2022; UNICEF, "Digital Age Assurance Tools and Children's Rights Online across the Globe" (2021); eSafety Commissioner Australia, Principles and background (2019).