Guidelines for industry on Child Online Protection

2020

You check your Amazon account (shared with your 15yo) and see an order for sex toys, robots and other items. What can YOU do?
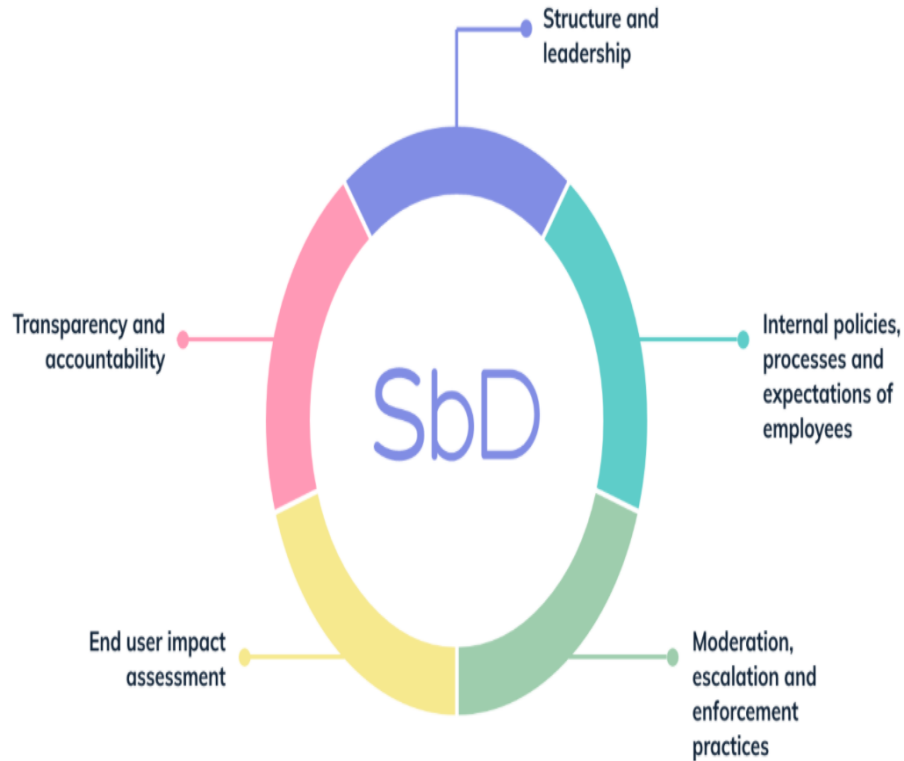
Safety By Design (SbD)

Safety by Design (SbD) principles that the eSafety Commissioner (eSafety), Australia

- The SbD underscores the importance of proactively considering user safety as a standard risk mitigation during the development process.
- Retrofitting safety is allowed however instead allowing users to experience harm beforeaddressing the problem SbD promotes anticipates risk at the design stage.
- SbD emphasizes the need to address online harms,
- alongside user safety and rights.

Safety By Design

Structure and leadership

Transparency and accountability

Internal policies, processes and expectations of employees

SbD

End user impact assessment

Moderation, escalation and enforcement practices

These principles provide guidance in incorporating, enhancing and assessing user safety considerations throughout the design, development and deployment phases of a typical service lifecycle.

The principles firmly place user safety as a fundamental design principle that needs to be embedded in the development of technological innovations from the start.

SbD Principles

At its heart are the SbD Principles, a model template for industry of all sizes and stages of maturity, providing guidance as they incorporate, assess and enhance user safety.

Principle 1: service provider responsibilities

The burden of safety should never fall solely upon the end user. Service providers can take preventative steps to ensure that their service is less likely to facilitate or encourage illegal and inappropriate behaviours.

Cont'

Principle 2: User empowerment and autonomy
The dignity of users is of central importance, with <u>users' best interests a primary consideration, through features, functionality</u> and an inclusive design approach that secures user empowerment and autonomy as part of the in-service experience.

Principle 3: Transparency and accountability
Transparency and accountability are critical to a robust approach to safety. <u>They not only provide assurances that services are operating according to their published safety objectives, but also assist in educating and empowering users about steps they can take to address safety concerns.</u>

Applying the Ethical Principles

P1
Ensuring that there are clear internal protocols for engaging with law enforcement, support services and illegal content hotlines.

P2
Providing technical measures and tools which easily allow users to manage their own safety while making sure procedures are set to the most secure privacy and safety levels by default.

P3
Embed user-safety considerations, training and practices into the roles, functions and working practices of all individuals who work with, for, or on behalf of the product or service.

Cont'

Develop community standards, terms of service and moderation procedures that are fairly and consistently implemented.

Provide built-in support functions and feedback loops for users that inform users on the status of their reports, what outcomes have been taken and offer an opportunity for appeal.

Carry out open engagement with a wide user base, including experts and key stakeholders, on the development, interpretation and application of safety standards and their effectiveness or appropriateness.

Cont'

Implement social contracts at the point of registration; these outline the duties and responsibilities of the service, user and third parties for the safety of all users.

Leverage the use of technical features to mitigate against risks and harms, which can be flagged to users at point of relevance, and which prompt and optimize safer interactions.

Commit to consistently innovate and invest in safety-enhancing technologies on an ongoing basis and collaborate and share with others safety-enhancing tools, best practices, processes and technologies.

Zoom Case: How the technology adheres to safety design principles

According to Zoom's April announcement, a number of measures were introduced by the company to address the security and privacy problems. Zoom announced a freeze in new feature development effective immediately and shifted all engineering resources to address security and privacy issues. The company conducted a comprehensive review with third-party experts and setup a CISO council.

Zoom has offered a series of training sessions, tutorials, free interactive daily webinars to users, and has taken steps to minimize support wait times so that customers could be empowered to use the various settings offered within the product to set up more secure meetings.

Such features that include:

– Limit meeting attendance to participants who are signed with their emails listed.

Zoom features Cont'

 Set up a waiting room function:
- – Password protect meeting access
- – Lock meetings once they start
- – Mute participants who are not presenting
- – Remove unwanted participants
- – Disable private chat

Zoom is preparing a transparency report to detail the information related to requests for data, records and content. The company's chief executive officer has hosted weekly webinars to answer questions from the community.

Zoom's Key challenges to Ethical design and privacy

- A series of security flaws and privacy breaches were found on Zoom.
- On 30 March 2020, the FBI issued a public warning about a problem named after Zoom – "zoom-bombing."
- The company was also said to have shared user data with third parties and make false claims about the encryption algorithm it uses. In June, Zoom shut down accounts of Chinese dissidents after an online Zoom meeting about the anniversary of 1989 Tiananmen Square social movement.
- On 1 April, Zoom published a blog post to respond to a list of criticisms and outlined a 90-day plan. On the 29th of June, the company appointed a new chief information security officer.

Same mistakes continue to expose children to online harms.

Self-regulation

Co-regulation

Retrofitting for safety is okay but what would work best in our peculiar situation?

Full regulation

A few hacks

- Your daughter was playing on your smartphone and someone Airdrops a photo that was offensive and inappropriate. What can YOU do?

- Does your child read on an e-reader? What do you like or dislike about the e-reader?

- Which is safer and easier for you as a parent to monitor for your 13-17 age category when you wish to limit exposure to certain friends?