

Placing user safety at the forefront of online service design

Australia's Office of the eSafety Commissioner commenced development of the Safety by Design (SbD) initiative in 2018, in recognition of the importance of proactively considering user safety as a standard risk mitigation during the development process, rather than retrofitting safety considerations after users have experienced online harm.

Safety by Design (SbD) is an initiative which places the safety and rights of users at the centre of the design, development and deployment of online products and services. Led in Australia by eSafety, the initiative aims to encourage and assist industry to take a proactive and consistent approach to user safety.

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. The Principles place user safety as a fundamental design consideration.



SbD 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, inflame or encourage illegal and inappropriate behaviours.

To help ensure that known and anticipated harms have been evaluated in the design and provision of an online service, a service should take the following steps:

1. Nominate individuals, or teams—and make them accountable—for user-safety policy creation, evaluation, implementation, operations.
2. Develop community standards, terms of service and moderation procedures that are fairly and consistently implemented.
3. Put in place infrastructure that supports internal and external triaging, clear escalation paths and reporting on all user-safety concerns, alongside readily accessible mechanisms for users to flag and report concerns and violations at the point that they occur.
4. Ensure there are clear internal protocols for engaging with law enforcement, support services and illegal content hotlines.
5. Put processes in place to detect, surface, flag and remove illegal and harmful conduct, contact and content with the aim of preventing harms before they occur.
6. Prepare documented risk management and impact assessments to assess and remediate any potential safety harms that could be enabled or facilitated by the product or service.
7. 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.
8. Consider security-by-design, privacy-by-design and user safety considerations which are balanced when securing the ongoing confidentiality, integrity and availability of personal data and information.

SbD Principle 2: User empowerment and autonomy

The dignity of users is of central importance, with users' best interests a primary consideration.

The following steps will go some way to ensure that users have the best chance at safe online interactions, through features, functionality and an inclusive design approach that secures user empowerment and autonomy as part of the in-service experience. Services should aim to:

1. Provide technical measures and tools that adequately allow users to manage their own safety, and that are set to the most secure privacy and safety levels by default.
2. Establish clear protocols and consequences for service violations that serve as meaningful deterrents and reflect the values and expectations of the user base.
3. 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 optimise safer interactions.
4. 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.
5. Evaluate all design and function features to ensure that risk factors for all users—particularly for those with distinct characteristics and capabilities—have been mitigated before products or features are released to the public.

SbD Principle 3: Transparency and accountability

Transparency and accountability are hallmarks of a robust approach to safety. 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.

To enhance users' trust, awareness and understanding of the role, and importance, of user safety:

1. 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.
2. Ensure that user-safety policies, terms and conditions, community standards and processes about user safety are visible, easy-to-find, regularly updated and easy to understand. Users should be periodically reminded of these policies and proactively notified of changes or updates through targeted in-service communications.
3. 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.
4. Publish an annual assessment of reported abuses on the service, alongside the open publication of meaningful analysis of metrics such as abuse data and reports, the effectiveness of moderation efforts and the extent to which community standards and terms of service are being satisfied through enforcement metrics.
5. 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.