Australian Government

Online safety charter

A statement of the Australian Government's expectations of online service providers

DECEMBER 2019

Contents

Preamble

Online service providers play an important, positive role in improving our citizens' quality of life and contributing to our economic and social development. They facilitate the connection with others, access to information, new ideas and entertainment, and opportunities for business and creation. They also provide revolutionary opportunities for individuals to express their opinions and can assist in holding those in power to account. The influence and power of online services in our lives is undeniable, but with that power comes a responsibility to society.

Governments around the world are facing the difficult challenge of protecting their citizens online. The proliferation of online abuse and threats targeting women, refugees and persons with disabilities, among others, are having significant and adverse impacts on social cohesion and national security. The sexual abuse of children, disinformation intended to undermine public discourse and democracy, and the incitement to support, assist and undertake terrorist acts, are equally issues of key concern.

The Australian Government recognises the immense challenges online service providers and governments alike face in protecting users online, including the scale of user activity, volume of user-generated content and current limits of technological solutions. Online service providers are not solely responsible for keeping Australians safe: online safety is a shared responsibility and all players must be accountable for creating a safer online world. This includes parents, carers, teachers and individuals, alongside governments and civil society.

However, as the conduits for access to the online environment, online service providers have a unique and important role to play in supporting safe online experiences.

There is a global conversation underway about what measures should reasonably be taken by technology companies to improve the online safety of users on their platforms. Leaders in the technology sector have invested in social responsibility programs, and voluntarily committed to domestic and international efforts to address harmful misuse of the internet, such as the Australian Taskforce to Combat Terrorist and Extreme Violent Material Online and the Christchurch Call to Action.

It is important that we are clear about what we — as a community — expect from industry in supporting online safety in Australia. That is why the Government has delivered on its commitment to develop an Australian Online Safety Charter. This Charter provides a high level articulation of the Australian Government's expectations of online service providers — to adopt responsible, proactive online safety measures in everything they do.

Only through preventative measures and a collective sense of responsibility will we be able to deliver the transformative changes required to make the online world a safer space.

Outline

Development

The Government's intention to develop an Online Safety Charter was announced on 16 December 2018 as part of the Keeping our Children Safe Online Package. A draft Charter was released for consultation on 16 February 2019 and 21 submissions were received in response from a range of stakeholders, representing the digital industry, non-government organisations, the community sector, state, territory and federal governments, the education sector, and academic experts. These valuable contributions have guided and informed the development of this document.

Purpose and status

This Charter articulates a set of community-led expectations for industry to protect citizens, especially children and vulnerable members of the community, from harmful online experiences. It is based on the premise that behaviour that is unacceptable offline should not be tolerated or enabled online. Online service providers have a responsibility to respect the rights and dignity of users online and to take meaningful action to address and prevent harms being incurred by those using their products or services.

A fundamental expectation of the Government is that digital services and products available to Australian users should have safety principles, protections and processes embedded into their design, development and deployment.

The Charter builds on the eSafety Commissioner's Safety by Design initiative, expanding and extending those expectations in areas of particular importance, such as age-appropriate content and support for law enforcement efforts.

The Charter sets out a number of additional expectations of Government in terms of empowering Australian users, upholding the integrity of their services, and collaborating with the Government and civil society. Importantly, the standards set out in the Charter are voluntary and there will be no sanctions for non-compliance. However, Government will take into account the extent to which technology firms and digital platforms operating in Australia are meeting the expectations set out in the Charter when assessing the need for further regulatory reform.

Scope and application

The Charter will have application to a range of firms operating in the Australian digital media market:

social media/social networking services

search engines

content hosts

gaming providers

internet service providers

app developers and app distributors

For the sake of simplicity, these operators are referred to as 'online services' in the Charter.

However, it is not intended that the Charter would have universal application to all such services. For many, the principles will simply not be relevant. For example, internet service providers do not moderate or publish content. For others, the application of these principles would be onerous and unmanageable. It would not be feasible, for example, for start-up firms to develop a common transparency standard for application across the industry, or to significantly invest in online safety tools such as content hashing.

It is also acknowledged that the digital media market is highly dynamic. New services and new platforms will continue to emerge over the coming years, while existing apps, services and products may wane in terms of their impact, influence and reach.

Moreover, the Charter is not intended to address other areas of concern with respect to the activities of technology firms and digital platforms, such as digital privacy, data portability, cyber security and impacts on competition. These issues will be considered by Government through other fora, for example in its response to the recommendations made by the Australian Competition and Consumer Commission in its Digital Platforms Inquiry Report released on 26 July 2019.

The Online
Safety Charter

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:

1.1 Nominate individuals, or teams — and make them accountable — for user-safety policy creation, evaluation, implementation and operations.

› Embed online safety considerations in the operating practices, governance arrangements and organisational culture of firms.

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

› Regularly review these standards, terms and procedures and revise them if there is evidence of systemic failure or persistent violations.

› Ensure that appeals processes are accessible to users who consider that they have been wrongly suspected or subject to other penalties incorrectly, and for the victims of abuse or harassment.

1.3 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.

› Assign the highest level of priority, in terms of content detection, triaging and moderation, to abhorrent content including child exploitation and sexual abuse material, terrorist and extreme violent content and content inciting suicide and self-harm.

› Provide guidance on how to refer complaints about age-inappropriate content, contact and conduct to regulatory authorities in Australia.

1.4 Ensure there are clear internal protocols for engaging with law enforcement, support services and illegal content hotlines.

› Provide Australian authorities with assistance, as reasonably required, for the purposes of law enforcement and safeguarding the national interest, and maximise automation in the flow of information to and from Australian authorities, ensuring that sensitive material is transmitted securely.

1.5 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.

› Where feasible and appropriate to the service, utilise technology to 'fingerprint' content that has been identified as illegal or harmful and deploy systems to prevent the attempted upload, re-upload or sharing of this material.

1.6 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.

1.7 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.

› Implement robust account validation processes, particularly for account creation, and implement systems that prevent the automated creation of accounts and the circumvention of account suspensions.

1.8 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.

› Where the service or product is likely to be accessed by children:

  › minimise the collection and disclosure of children's personal data and avoid its detrimental use;

  › uphold rules and behaviour standards, including age restrictions; and

  › provide accessible reporting tools and guidance for parental controls.

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:

2.1 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.

› Implement mechanisms, including parental controls and access controls that effectively protect children from online harms, such as grooming, exposure to sexually explicit or graphically violent content, abuse, threats or cyberbullying.

› Make information about online safety available at all relevant points in the supply chain, including point-of-purchase, registration, account creation, first use and when they are upgraded or updated.

› Invest in tools and features that provide users with control over what they share, who they share with, the content they see or experience, and who can contact them.

› Actively promote age-appropriate online safety resources to Australian users and signpost support services that are relevant to Australian users.

2.2 Establish clear protocols and consequences for service violations that serve as meaningful deterrents and reflect the values and expectations of the user base.

› Promote positive and respectful user behaviour and engagement and make clear the user's legal responsibilities in relation to the content they upload or share.

› Ensure that the minimum standards for behaviour and engagement are communicated to users.

› Take decisive action in instances of the violation of policies, terms of use and standards.

2.3 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.

› Where appropriate and feasible for the service, provide warnings, content labelling and age ratings/guidance containing sufficient information about the nature of high impact content to allow for viewer discretion.

› Where appropriate and feasible for the service, invest in robust age-verification systems and restricted access systems that enforce age limits and prevent users under 18 years from accessing harmful content, including age-inappropriate content.

2.4 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.

2.5 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.

› Build sufficient flexibility into the design of a product or service so that service providers have the capacity and capability to effectively address the safety of the service as it expands.

› For firms along the supply chain, such as app and game distributers, only accept products that adhere to the above safety-by-design principles.
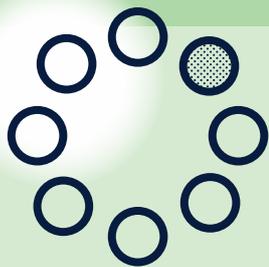
3 Transparency and accountability

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:

3.1 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.

3.2 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.

› Support community programs and initiatives that promote resilience and respectful online engagement.

3.3 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.

› Collaborate with public researchers and civil society to better understand users' safety on their services and the processes to ensure this.

3.4 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.

› Develop a common transparency reporting standard with like services and cross platform reporting protocols for online safety matters of global concern, including child sexual abuse material and terrorist and extreme violent material.

3.5 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.

› Support the capacity of smaller service providers to improve the safety of their services through the sharing of best practice and best technological solutions.

**Australian Government
Online safety charter**

December 2019