January 22, 2024

Submission from Apple Pty Limited ("Apple")
on Draft Online Safety Industry Standards[1]

Introduction

At Apple, our goal has been and always will be to create technology that empowers and enriches people's lives—while helping them stay safe. Our work to protect users has never been more important, as the myriad threats they face in the modern digital world continue to grow in complexity and frequency each and every year. Simultaneously, users continue to increase the data they store on their devices and in the cloud, both the amount and sensitivity. Every day, users rely on their devices and technology services to securely store and process their health data, the location of their family members, their intimate messages, and financial data, among other things. Our users entrust us with keeping this data safe, secure and private.

Apple fully shares concerns around the proliferation of abhorrent child sexual abuse material ("CSAM") and pro-terror content. We are committed to keeping our users safe and we will continue to invest in technologies that do just that. Yet, we have serious concerns that the draft standards pose grave risks to the privacy and security of our users and set a dangerous global precedent.

eSafety's mandate is to "promote online safety for all Australians."[2] This makes sense; from our perspective and that of many of the world's leading technologists, the promise of online safety is only made possible when strong and robust security measures exist to protect the information of law-abiding citizens against increasingly aggressive data thieves, sophisticated state-sponsored threat actors, and other malicious attackers.

Rather than making all Australians safer, the draft standards risk undermining fundamental privacy and security protections. As currently drafted, eSafety could seek to require providers to: (1) build backdoors into end-to-end encrypted services to monitor data that providers are currently unable to access, weakening the most effective security protection users have to safeguard their data; (2) scan the private communications and files of every Australian user; (3) comply with standards that are inconsistent with the laws and regulations in other countries in which providers operate;

[1] This submission covers both the Designated Internet Services Standard (Draft Online Safety (Designated Internet Services – Class 1A and Class 1B Material) Industry Standard 2024, eSafety Commissioner (Nov. 20, 2023)) ("DIS Industry Standard"), and the Relevant Electronic Services Standard (Draft Online Safety (Relevant Electronic Services – Class 1A and Class 1B Material) Industry Standard 2024, eSafety Commissioner (Nov. 20, 2023)) ("RES Industry Standard"). Collectively, these standards are referred to throughout as "draft industry standards" or "draft standards."

[2] *eSafety Commissioner*, Department of Infrastructure, Transport, Regional Development, Communications and the Arts, https://www.infrastructure.gov.au/media-technology-communications/internet/online-safety/esafety-commissioner (last visited Dec. 19, 2023) (emphasis added).

and (4) serve as agents of the state, surveilling user data in a way that even law enforcement today cannot legally require.

We believe there are alternative ways to achieve the goal of combatting abhorrent content that do not require undermining the privacy and security of all Australians and we urge eSafety to allow providers flexibility to pursue those means.

The Draft Standards Could Require Backdoors into End-to-End Encryption

End-to-End Encryption is one of the Most Effective Security Technologies Available

End-to-end encryption ensures that only users—and not the companies who provide services—can access a user's personal information and communications. Encryption provides an essential layer of additional security because it ensures that a malicious actor cannot obtain access to a user's data even if the actor is able to breach a service provider's networks. It shields everyday citizens from unlawful surveillance, identity theft, fraud, and data breaches, and it serves as an invaluable protection for journalists, human rights activists, and government employees who are constantly targeted by malicious actors. The critical value provided by encryption—and end-to-end encryption in particular—is a key reason for the global technology community's broad consensus in support of these features.

Encryption protects every single Australian, including children, and is essential for preserving the collective public safety and national security. The devices that Australians carry contain their personal messages, health information, and photos—and often, parents' devices contain their child's sensitive health information and personal photos as well. Devices Australians carry are also conduits to businesses, infrastructure, and other critical services. Vital infrastructure—such as power grids and transportation hubs—becomes more vulnerable when connected devices get hacked. Criminals, terrorists, and state-sponsored actors who want to infiltrate systems and disrupt sensitive networks may start their attacks by accessing just one person's smartphone. In the face of these threats, now is not the time to weaken encryption. There is a profound risk of making the jobs of malicious cyber actors easier, not harder. And increasingly stronger—not weaker—encryption is the best way to protect against these threats.

The need for strong cyber protections like end-to-end encryption is clear. According to the Deputy Prime Minister and Minister for Defence, the Hon Richard Marles MP, "[r]ecent global and national events have demonstrated the growing threat to Australia by malicious cyber actors."[3] Threats to user data are undeniably growing. The total number of data breaches more than tripled between 2013 and 2022—exposing 2.6 billion personal records in the past two years alone—and has continued to get worse in 2023. According to Reuters, "nearly half [Australia's] 26 million population had personal

[3] *Release of the annual Cyber Threat Report 2022-23*, Australian Government Defence Ministers (Nov. 15, 2023), https://www.minister.defence.gov.au/media-releases/2023-11-15/release-annual-cyber-threat-report-2022-23.

information stolen in just two data breaches at companies, while a cyber attack at its biggest port operator this month brought supply chains to a standstill."[4] The Australian Signals Directorate recently found that the average cost of cybercrime increased by 14 percent and the number of cybercrime reports jumped by 23 percent in the last year.[5]

Apple has continued to enhance its security and safety features over time because the threats to users and their data are relentless, pervasive, sophisticated, and evolving. Our customers around the world expect us to protect them and their personal information from bad actors who seek to access, steal, and use it without a user's permission. Apple and other providers, in line with best practices, rely on end-to-end encryption as a critical method to do so.

The Draft Standards Lack Meaningful Protections for End-to-End Encryption

We appreciate that eSafety has indicated its intent to preserve end-to-end encryption. For example, it has stated that providers are not expected to "design systematic vulnerabilities or weaknesses into end-to-end encrypted services."[6] That sentiment, however, is not explicitly stated anywhere in the actual draft RES and DIS standards. And, as we explain in more detail below, while the draft standards have some limits based on technical feasibility, those provisions cannot be read as adequately supporting end-to-end encryption protections.

The approach in the draft standards stands in stark contrast to the clear protection for end-to-end encryption that eSafety accepted and registered in prior industry codes. The Head Terms for the Consolidated Industry Codes of Practice for the Online Industry state that the codes do not require:

> "[A]ny industry participant to undertake steps that do the following:
>
> (a) implement or build a systematic weakness, or a systematic vulnerability, into a form of encrypted service or other information security measure;
>
> (b) build a new decryption capability in relation to encrypted services;
>
> (c) render methods of encryption less effective;

[4] Byron Kaye, *Australia beefs up cyber defences after major breaches*, Reuters (Nov. 22, 2023), https://www.reuters.com/technology/cybersecurity/australia-goes-cyber-offensive-with-sweeping-resilience-plan-2023-11-22.

[5] *ASD Cyber Threat Report 2022-2023*, Australian Signals Directorate (Nov. 14, 2023), https://www.cyber.gov.au/about-us/reports-and-statistics/asd-cyber-threat-report-july-2022-june-2023.

[6] *See Discussion Paper: Draft Online Safety (Relevant Electronic Services - Class 1A and 1B Material) Industry Standard 2024 and Draft Online Safety (Designated Internet Services - Class 1A and 1B Material) Industry Standard 2024*, eSafety Commissioner, at 14 (Nov. 2023), https://www.esafety.gov.au/sites/default/files/2023-11/Discussion-Paper-draft-Online-Safety-Standards-(Class-1A-and-1B).pdf.

> (d) undertake monitoring of private communications between end-users."[7]

eSafety claims the same protections for end-to-end encryption in the codes apply to the standards but this is not supported by any language to that effect. We recommend that eSafety adopt a clear and consistent approach expressly supporting end-to-end encryption so that there is no uncertainty and confusion or potential inconsistency across codes and standards.

Similarly, eSafety asserts that providers will not be required to build any systems, processes or technologies when doing so is not technically feasible. However, eSafety appears to define technical feasibility narrowly, focusing primarily on cost at the expense of other important factors. Specifically, eSafety states that technical feasibility should be assessed primarily based on the "expected financial cost to the provider of taking the action" and "whether it is reasonable to expect the provider to incur the cost."[8] Focusing on financial cost at the expense of other important considerations creates an imbalance harmful to broader societal interests. For instance, such a narrow approach to technical feasibility does not take into account how technology companies appropriately consider whether a particular product design change is in the best interests of securing its users.

eSafety's seemingly narrow view of technical feasibility is an outlier, departing dramatically from other similar regimes. For instance, the Australian Assistance and Access Act requires that technical measures be reasonable, proportionate, practical and technically feasible and must include consideration of "industry interests, necessity, privacy, cyber security and intrusiveness."[9] Outside of Australia, the UK's online safety regulator recently promulgated similar guidance, clarifying that their view of technical feasibility includes an assessment of whether required changes would materially compromise the security of a service.[10]

[7] *See Consolidated Industry Codes of Practice for the Online Industry (Class 1A and Class 1B Material) Head Terms*, eSafety Commissioner, at Head Term 6.1 (Sept. 12, 2023), https://www.esafety.gov.au/sites/default/files/2023-09/Consolidated-Industry-Codes-of-Practice-Head-Terms-12-September-23.pdf.

[8] *See* DIS Industry Standard at Sec. 7; RES Industry Standard at Sec. 7.

[9] *Assistance and Access: Industry assistance framework - limitations and safeguards*, Department of Home Affairs, https://www.homeaffairs.gov.au/about-us/our-portfolios/national-security/lawful-access-telecommunications/assistance-and-access-limitations-safeguards (last visited Nov. 30, 2023); *see also* Investigatory Powers Act § 255(4), https://www.legislation.gov.uk/ukpga/2016/25/section/255/enacted (requiring separate review of cost and technical feasibility).

[10] *See Protecting people from illegal harms online*, Ofcom, at Section 14.16 (Nov. 9, 2023), https://www.ofcom.org.uk/__data/assets/pdf_file/0022/271147/volume-4-illegal-harms-consultation.pdf ("However, our measures would not apply to services that are technically unable to analyse user-generated content present or disseminated on the service to assess whether it is content of a particular kind, particularly where such changes as would need to be made to enable this would materially compromise the security of the service.").

We urge eSafety to revise the definition of technical feasibility so that it more closely aligns with other approaches in and outside of Australia that explicitly take into account risks to user security and privacy, among other factors.

Automated Content Scanning Creates Unintended Privacy and Security Risks

Child sexual abuse material is abhorrent and we are committed to breaking the chain of coercion and influence that makes children susceptible to it. Equally, we firmly believe that the interests of Australians will be severely harmed if eSafety compels providers to comb through their users' most sensitive personal information—which Apple does not do. We do not take this position lightly. It is one that we have reached after years of investment, research and development, thoughtful consideration, and public engagement. We worked hard to conceptualize a hybrid device-server technology to detect known CSAM in iCloud Photos without compromising privacy and security. Ultimately, after having consulted extensively with child safety advocates, human rights organizations, privacy and security technologists, and academics, and having considered scanning technology from virtually every angle, we concluded that it was not practically possible to implement without ultimately imperiling the security and privacy of our users.

The Serious Risks of Scanning Users' Personal Information

Some companies regularly scan personal information in the cloud to monetize the information of their users. Apple does not. We have chosen a very different path—one that prioritizes the security and privacy of our users. As threats become increasingly sophisticated, we are committed to providing consumers with the best data security in the world, and we constantly identify and mitigate emerging threats to users' personal information, on devices and in the cloud. Scanning every user's privately stored iCloud data would pose serious security and privacy problems. Among other things, it could create new threat vectors for data thieves to find and exploit.

Scanning for particular content opens the door for bulk surveillance of communications and storage systems that hold data pertaining to the most private affairs of many Australians. Such capabilities, history shows, will inevitably expand to other content types (such as images, videos, text, or audio) and content categories. A tool for one type of surveillance can be reconfigured to surveil for other content, such as that related to a person's familial, political, professional, religious, health, sexual, reproductive, or other activities. Tools of mass surveillance have widespread negative implications for freedom of opinion and expression and, by extension, democracy as a whole. For example, awareness that the government may compel a provider to watch what people are doing raises the serious risk of chilling legitimate associational, expressive, political freedoms, and economic activity.[11]

[11] *See The Economic Impact of Laws that Weaken Encryption*, Law & Economic Consulting Associates (April 5, 2021), https://www.internetsociety.org/resources/doc/2021/the-economic-impact-of-laws-that-weaken-encryption/.

Scanning all privately stored and communicated data can also lead to the circumvention of legal processes—which is inconsistent with the objective of providing appropriate community safeguards. Our law enforcement team works tirelessly to provide law enforcement agencies, including Australia's, with data we have that is responsive to lawful requests related to CSAM and terrorism. This level of commitment reflects Apple's view of the abhorrent nature of such content. Those requests must be supported by lawful investigations with demonstrable cause. Forcing providers to comb through the private storage and communications of all its users, without any particularity, reason for suspicion, or other constraint, improperly turns private companies into arms of the state and would upend the trusted relationship between a provider and its users.

The risks this technology poses are not limited to Australians; designing this technology for one government could spur other countries to follow suit. The standards set in Australia will have far-reaching global repercussions; countries that lack the robust legal protections afforded to Australians will leverage this and expand on it. Surveillance ordered by governments unconstrained by civil rights and civil liberties protections harm the internationally recognized and fundamental human rights of people whose data would be collected by network scanning, especially opposition politicians, journalists, activists, and members of marginalized groups. If such governments know that service providers have put into place scanning systems pursuant to mandates from the Australian government, they will seek to use those systems for their own purposes: if we are forced to build it, they will come. So too will criminal actors, drawn to where protections are weaker and preying on innocent users is easier.

The Overly Broad and Vague Proposed Definitions Harm Legitimate Free Expression Without Advancing eSafety's Anti-CSAM and Anti-Terror Content Goals

We share the goal of making the Internet a safer place for children and all users. We are focused on developing technologies that empower children and their parents to combat these harms. eSafety's requirement to surveil user's personal information perpetuates the misconception that content detection technologies are a panacea. However, scanning systems are not foolproof. There is evidence from other platforms that innocent parties have been swept into dystopian dragnets that have made them victims when they have done nothing more than share perfectly normal and appropriate pictures of their babies.[12]

This concern is amplified by the overly broad and vague definitions that the draft standards use for certain key terms, some of which are highly subjective and susceptible to multiple interpretations. Such terms include, for example, "child sexual abuse material," "child sexual exploitation material," and the definitions related to "crime and violence material," "drug-related material," "extreme crime and violence material," and "pro-terror material."

[12] *See How Your Child's Online Mistake Can Ruin Your Digital Life*, The New York Times (Nov. 27, 2023), https://www.nytimes.com/2023/11/27/technology/google-youtube-abuse-mistake.html.

For example, the critical term “child sexual abuse material” is not limited to CSAM images or videos. Instead, it includes “material that . . . describes . . . child sexual abuse.”[13] Presumably, this would include everything from memoirs, and “Me Too” statements that discuss a victim’s experience with child sexual abuse to news reports and government criminal charging documents that explain how a defendant is alleged to have abused a child. The term “child sexual exploitation material” is also defined exceedingly broadly and vaguely to include certain material that is “exploitative,” “offensive,” “gratuitous,” or “likely to cause offense to a reasonable adult.”[14] Those vague terms are highly subjective and give little guidance to providers who, under the current draft standards, would have to find and remove such material.

As another example, the definition of “pro-terror material” includes a potentially broad carve out that is similarly ill-defined: “[h]owever, material accessible using a [relevant electronic service or designated internet service] is not pro-terror material if its availability on the service can reasonably be taken to be part of public discussion, public debate, entertainment or satire.”[15] Assessing whether people praising a terrorist act are doing so in violation of the standards or as part of a “public discussion” or “public debate” would be exceedingly difficult for providers to do.

In connection with these definitional issues, we are concerned that, out of an abundance of caution and in order to avoid running afoul of the requirements of the standards if implemented as currently drafted, providers will have no practical choice but to remove significantly more lawful content than strictly necessary if the definitions were more narrowly and appropriately focused. Such unnecessary removals, while the rational response to liabilities sought to be imposed by the draft standards, harm the legitimate free expression and political communication rights of Australians without advancing eSafety’s anti-CSAM and anti-terror content goals.

The Draft Standards Could Conscript Providers to Surveil Users in a Way that Even Law Enforcement Cannot Require

The draft standards adopt a fundamentally different approach to “search and seizure” of private communications and material than that which applies to law enforcement and other government agencies. The draft standards can be construed to require providers to take actions that law enforcement could not otherwise take under Australian law. In some instances, for example, the law requires the government to obtain a warrant or other court order to search and seize private communications and data. Yet, the draft standards require no such due process, compelling providers to indiscriminately search vast streams of protected communications and large stores of personal information, and, in some cases, seize content based on ill-defined and vague definitions.

[13] *See* DIS Industry Standard at Sec. 6(1); RES Industry Standard at Sec. 6(1).

[14] *See* DIS Industry Standard at Sec. 6(1); RES Industry Standard at Sec. 6(1).

[15] *See* DIS Industry Standard at Sec. 6(1); RES Industry Standard at Sec. 6(1).

This approach could upend the balance struck by years of Australian jurisprudence, transforming providers into agents of the government. Such a fundamental shift should be the subject of an Act of Parliament, not subordinate legislation, with appropriate protection for providers of services required to undertake such actions (such as would apply under other regimes requiring providers to assist law enforcement, for example, section 317ZJ of the Telecommunications Act). There is a risk that providers could face civil liability and legal costs for actions taken at the behest of a government agency. There is no provision in the Online Safety Act or the draft standards that grants immunity to providers for good faith actions that they take pursuant to eSafety's standards or provides that the government will indemnify them for such actions.

Apple's Approach

Apple devices, such as Macs, iPhones, iPads, and Apple Watches, play an important role in the lives of many individuals and families across the world. And for years, Apple has built innovative features to protect its users.[16] For example, in December 2021, we introduced Communication Safety, a feature designed to address online exploitation and the unwanted sharing of certain images.[17] Communication Safety aims to help keep children safe by intervening and giving a child a moment to pause when they receive or attempt to send images that contain nudity.[18] The feature provides guidance and age-appropriate resources to help them make a safe choice, including the choice to contact someone they trust. The goal is to disrupt the grooming of children by making it harder for predators to normalize this behavior and to create a moment for a child to think when facing a critical choice.

We continue to enhance Communication Safety by expanding the feature to more easily and more broadly protect children. First, the feature is now on by default for accounts of all children under thirteen years of age and available worldwide. Second, the feature covers both video content and still images. And we have expanded these protections to more areas across the Apple ecosystem, including AirDrop, the system Photo Picker, FaceTime messages, and Contact Posters in the Phone app.

Children, of course, use third-party apps in which they can share content, and app developers also want to protect children in vulnerable situations. As a result, we brought Communication Safety to the entire platform with the Sensitive Content Analysis framework. This framework enables developers to detect sensitive content in their apps

[16] *See, e.g.*, *Use parental controls on your child's iPhone, iPad, and iPod touch*, Apple (Nov. 1, 2023), https://support.apple.com/en-us/HT201304.

[17] *See Expanded Protections for Children*, Apple, https://www.apple.com/child-safety/ (last visited Nov. 30, 2023).

[18] *See About Communication Safety on your child's Apple device*, Apple (Oct. 25, 2023), https://support.apple.com/en-us/HT212850.

for people who have enabled Communication Safety or Sensitive Content Warning,[19] and, due to its ease of implementation, developers of communication apps have incorporated this advanced technology into their products. Importantly, these features use privacy-preserving technology – all image and video processing occurs on device, meaning that Apple does not get access to the content.

Apple has also expanded guidance in Siri, Spotlight, and Safari Search by providing additional resources to help children and parents stay safe online and obtain assistance if they encounter unsafe situations.[20] For example, users who ask Siri how they can report child exploitation will be pointed to resources for where and how to file a report. Siri, Spotlight, and Safari Search also have been updated to intervene when users perform searches for queries related to child exploitation. These interventions explain to users that interest in this topic is dangerous and illegal and provide resources from partners who can provide help to prevent abuse.

Apple also provides a number of other features and applications that empower parents in managing their children's device usage, many of which are accessible through Apple's Screen Time feature. Screen Time gives parents a better understanding of the time that their kids spend using Apple devices, and it allows parents to set limits on specific apps and features.[21] Users can easily turn on Screen Time from the Settings menu of their child's device and set a passcode to lock in protective settings, such as extending the app time limit duration. Once activated, Screen Time will provide a report about how the device has been used, apps that have been opened, and websites that have been visited. All of this information can be very helpful for a parent who wishes to safeguard their child's activity on an Apple device and make sure that it is being used in a safe and responsible way.

Apple also empowers parents to utilize Screen Time to control a number of other crucial safety functions. Our Communication Limits feature allows parents to define specific contacts with whom their child may communicate. And Content Restrictions let parents restrict their child's ability to listen to music with explicit content, access adult websites, or watch movies or TV shows with specific ratings.[22] Apps also have ratings so that parents can limit their children to appropriate apps using Content Restrictions. Screen

[19] *See SensitiveContentAnalysis*, Apple Developer, https://developer.apple.com/documentation/sensitivecontentanalysis (last visited Nov. 30, 2023);
see also Detecting nudity in media and providing intervention options, Apple Developer, https://developer.apple.com/documentation/sensitivecontentanalysis/detecting-nudity-in-media-and-providing-intervention-options (last visited Nov. 30, 2023).

[20] *See Expanded Protections for Children*, Apple, https://www.apple.com/child-safety/ (last visited Nov. 30, 2023).

[21] *See Use Screen Time on your iPhone, iPad, or iPod touch*, Apple (Sept. 12, 2023), https://support.apple.com/en-us/HT208982.

[22] *See Use parental controls on your child's iPhone, iPad, and iPod touch*, Apple (Nov. 1, 2023), https://support.apple.com/en-us/HT201304.

Time also allows parents to prevent certain web content and restrict Game Center. Web Content Restrictions help prevent children from accessing adult websites while they browse the web. In iOS 17, adult websites will be blocked on child devices and parents can choose to allow or block additional websites. There's also an option for parents to block all websites, except for a specific set that they choose to allow for their children. This tool helps parents play a direct role in keeping their kids safe online, while allowing children to have access to valuable educational and entertainment resources.

Our approach seeks to tackle the issue of child online sexual exploitation without raising the serious privacy and security concerns that come with compelled surveillance of each and every Australian user's privately stored personal information and private communications. We continue to innovate and develop new features and protections for children, and will continue to work collaboratively with child safety organisations, technologists, and government stakeholders like yourself on solutions that help protect the most vulnerable members of our society. Apple will continue investing in technologies to protect our users because it is the right thing to do, and we urge eSafety to preserve the strength of those technologies and the protection that they offer users.

Conclusion

Apple has and will continue to demonstrate its commitment to fighting the proliferation of abhorrent CSAM and pro-terror content. We will continue to work hard to make the Internet safer for Australians and our users everywhere. We appreciate eSafety's work on the draft standards and look forward to working together to implement a standard that supports the best outcomes for all users. We remain deeply concerned that some aspects of the draft standards do not sufficiently take into consideration the potential impacts on users, including in relation to their privacy and data security. We hope that our concerns about the draft standards' scanning requirement and its treatment of end-to-end encryption and technical feasibility will be addressed in the final versions.