COMMUNICATIONS ALLIANCE LTD

Communications Alliance Submission

to the Department of Infrastructure, Transport, Regional Development, Communications and the Arts

Statutory Review of the Online Safety Act 2021
Issues Paper

5 July 2024

CONTENTS

COMMUNICATIONS ALLIANCE		2
1.	INTRODUCTION	3
2.	CONTEXT OF THE STATUTORY REVIEW OF THE ONLINE SAFETY ACT 2021	5
3.	OVERARCHING COMMENTS	6
	Delegation to subordinate instruments and interpretation by the regulator	7
	Risk-based, proportionate approach	8
4.	SERVICES / EQUIPMENT IN SCOPE OF THE ACT	10
	Relevant electronic services	11
	Texts and MMS	11
	Designated internet services	13
	Websites and a multitude of other services	*13*
	Streaming video on demand services (SVODS)	*14*
	Internet carriage services	15
5.	MATERIALS AND BEHAVIOURS IN SCOPE OF ACT	15
	Interaction of the Act and the National Classification Scheme	15
	Pornography and children's access to age-inappropriate content	16
	Expansion to specific types of class 2 material	17
	Hate speech	19
6.	DUTY OF CARE	20
7.	BEST INTERESTS OF THE CHILD	21
	Interaction with other processes addressing children's rights online	21
	Relationship of best interests of the child, statutory duty of care and safety-by-design	21
	Reference framework for the best interests of the child	22
	Services in scope for best interests of the child considerations	22
8.	CODES/STANDARDS & ENFORCEMENT	23
9.	PENALTIES	24
10.	OTHER ISSUES	26
	Operation of sections 232 and 235 of the OSA	26
	Limitations of liability for voluntary action	26
ANNEX 1 – DEFINITIONS OF PORNOGRAPHY		28

Communications Alliance

Communications Alliance is the primary communications industry body in Australia. Its membership is drawn from a wide cross-section of the communications industry, including carriers, carriage and internet service providers, content providers, platform providers, equipment vendors, IT companies, consultants and business groups.

Its vision is to be the most influential association in Australian communications, co-operatively initiating programs that promote sustainable industry development, innovation and growth, while generating positive outcomes for customers and society.

The prime mission of Communications Alliance is to create a co-operative stakeholder environment that allows the industry to take the lead on initiatives which grow the Australian communications industry, enhance the connectivity of all Australians and foster the highest standards of business behaviour.

For more details about Communications Alliance, see https://www.commsalliance.com.au .

1. Introduction

1.1. Communications Alliance welcomes the opportunity to make a submission to the Department of Infrastructure, Transport, Regional Development, Communications and the Arts (Department) in response to the Statutory Review of the *Online Safety Act 2021 Issues Paper* (Issues Paper).

1.2. Communications Alliance and its members take very seriously their roles in relation to the safety of users online. All members invest substantial amounts of resources and time into systems, processes and/or technologies that aim to reduce material or activity on their products or services that is unlawful or harmful.

1.3. We welcome initiatives that aim to further online safety through a practical and holistic framework that aligns to and harmonises with other legislative and regulatory frameworks. We and our members are keen to constructively engage with all stakeholders, including the Department, the Office of the eSafety Commissioner (eSafety) and other industry sectors, on approaches that further improve the safety of Australians online, particularly those that are most vulnerable.

1.4. We have not sought to address all or specific questions in our submission but instead provide feedback on issues relevant to our members in relation to many of the consultation questions.

1.5. We also note that our feedback in response to the exposure draft of the *Online Safety (Basic Online Safety Expectations) Amendment Determination* (February 2024) is, in our view, still valid, and should be read in conjunction with this submission.

1.6. The *Online Safety (Designated Internet Services – Class 1A and Class 1B Material) Industry Standard 2024* and the *Online Safety (Relevant Electronic Services – Class 1A and Class 1B Material) Industry Standard 2024* were registered on 21 June 2024, i.e. in last stages of developing this submission. Therefore, we were unable to assess the detail of the registered standards, and our feedback on the Issues Paper is based on the draft standards published in January 2024.

1.7. Communications Alliance members may also make individual submissions.

1.8. The submission provides feedback in response to a multitude of issues. Those include:

- Concern with the concurrence of a multitude of interrelated processes, including in the areas of online safety, privacy, digital identity, AI, and social media and digital platforms in Australia more broadly;
- The need to ensure the review considers the long term efficacy, flexibility, complementarity and other factors, not only of the Online Safety Act (OSA) but also of the broader regulatory framework for online safety and other adjacent regulatory frameworks, including their alignments with international approaches;
- Concern with the persistent and substantial delegation of key concepts to subordinate legislation and/or discretionary interpretative powers of the regulator, thereby substantially limiting parliamentary debate and unduly allowing discretionary interpretation and enforcement of regulation;
- Concern with the lack of a risk-based and proportionate approach to the OSA and Basic Online Safety Expectations (BOSE) more broadly, and specifically within a number of definitions for online sections, including relevant electronic services, designated internet services and internet carriage services;
- Alternative considerations if a duty of care was indeed contemplated to be enshrined into the OSA, cautioning against importing new regulatory mechanisms from other jurisdictions without full consideration of the impact and effectiveness of the broad range of existing obligations;

- Concern with the application of the National Classification Scheme (NCS) to the online environment, particularly at scale in relation to user-generated content, and the resulting differential treatment of and outcomes for professionally produced content;
- The need to apply a consistent definition to pornographic content that recognises the different harms emanating from different media and types of pornographic representations;
- The need to ensure that any proposal relating to age-verification/assurance ought to strike an appropriate balance of efficiency, effectiveness and the impact/risk of harm vectors, given the potential privacy and cyber-security concerns with the collection of personal information;
- Querying the need for and/or concerns with the proposals to expand the types of behaviour and materials regulated under the OSA without appropriate parliamentary debate in relation to these specific types of materials, and whether sector-specific regulation is the appropriate means to address any potential issues, including in relation but not limited to hate speech;
- A number of considerations in relation to a potential 'best interests of the child' approach, including the need to appropriately target such considerations at services that are directed at children, rather than merely being accessible by children;
- The need to ensure that the Online Content Scheme of the OSA allows for an effective and efficient development of industry codes, e.g. through the ability to develop codes for individua sub-sections of the online sector and appropriate timeframes for code development;
- The need to consider the broader legislative framework for online safety and adjacent regulatory frameworks (including privacy legislation), and existing powers granted to eSafety (including those that have the ability to lead to significant reputational damage) when comparing penalty regimes;
- The need to clarify the interaction between the OSA and the *Telecommunications (Interception and Access) Act 1979* (TIA Act), to ensure the consistent application of the principle of privacy of communications as enshrined in the TIA Act; and
- The need to expressly exclude liability for actions taken voluntarily or under subordinate regulation by service providers.

For questions in relation to this submission, please contact Christiane Gillespie-Jones [redacted].

2. Context of the statutory review of the Online Safety Act 2021

2.1. The statutory review (review) of the *Online Safety Act 2021* (OSA) occurs against the background of no fewer than fifteen other reform processes. Those include:

1. the recently made *Online Safety (Basic Online Safety Expectations*) Amendment Determination 2024 [we will refer to the Amendment together with the *Online Safety (Basic Online Safety Expectations) Determination 2022* as BOSE];
2. the making of industry standards for class 1A and class 1B material for relevant electronic services (RES) and designated internet services (DIS) by the Office of the eSafety Commissioner (eSafety) (Phase 1 Standards);
3. the recent Government announcement of a pilot of age assurance technology;
4. the second stage of the modernisation of Australia's National Classification Scheme (NCS);
5. the development of industry codes for class 1C and class 2 material (Phase 2 Codes);
6. processes for the making of necessary subordinate regulation (rules/standards) as a consequence of the recently passed *Digital ID Bill 2024*, together with the *Digital ID (Transitional and Consequential Provisions) Bill 2024*;
7. the review of the *Privacy Act 1988* (Privacy Act) (with a foreshadowed introduction into Parliament in August 2024), including Government's agreement to implement a *Children's Online Privacy Code* to promote the design of certain services in the 'best interests of the child' and to introduce new provisions to address doxxing activity;[1]
8. foreshadowed legislation addressing hate speech and religious discrimination;[2]
9. the *Misinformation and Disinformation Bill 2023* and associated processes;
10. the voluntary code for online dating services;
11. the recently established Joint Select Committee on Social Media and Australian Society;
12. the establishment of the Select Committee on Adopting Artificial Intelligence, which will inquire into, among other things, the risks and harms arising from the adoption of AI technologies, and emerging international approaches to mitigate AI risks;
13. activity flowing from Government's interim response to the *Safe and responsible AI in Australia* consultation, including the proposed AI Safety Standard to be co-designed with industry and potential mandatory requirements for high-risk use cases;
14. the anticipated Government response in relation to dispute and complaints resolution processes of digital platforms, flowing from the ACCC's *Digital Platform Inquiry*; and
15. the Department of Home Affairs report in relation to understanding algorithms on digital platforms.

2.2. All of the above processes either directly influence the review or have the potential to substantially influence operational and design aspects of the services in scope of the

[1] p. 15, Australian Government, *Government Response Privacy Act Review Report*, 28 Sept 2023, as accessed at https://www.ag.gov.au/rights-and-protections/publications/government-response-privacy-act-review-report on 6 Feb 2024: "To clarify how the best interests of the child should be upheld in the design of online services, and provide further guidance on how entities are expected to meet requirements regarding targeting, direct marketing and trading, the Government **agrees** a Children's Online Privacy code should be developed (*proposal 16.5*) as soon as legislated protections for children are enacted to enable the development of such an APP code. The code would apply to online services that are likely to be accessed by children. To the extent possible, the scope of the code should align with international approaches, including the UK Age Appropriate Design Code, with similar exemptions for particular entities such as counselling services. The code developer should consult broadly with children, parents, child development experts, child welfare advocates and industry in developing the code."

[2] Media Conference – Parliament House, 13 Feb 2024, https://ministers.ag.gov.au/media-centre/transcripts/media-conference-parliament-house-13-02-2024 as accessed on 21 May 2024

OSA and the material subject to the provisions of the OSA. Conversely, the OSA itself forms the basis for a number of the processes above.

2.3. Consequently, we reiterate our concern with and bewilderment at the concurrence of the above processes. We hold serious concerns that inconsistent, unnecessarily complex and/or impractical legislation/regulation may arise as a result of the multitude of reform processes being undertaken in parallel. There are key policy decisions regarding the intersection of these different legislative regimes (which are often overseen by different regulators) that are yet to be resolved. Running these various reform processes in parallel risks continuing or increasing inconsistencies and issues that arise from these intersections. It is, in our view, also unrealistic to expect our sector to provide optimal feedback into these processes and to implement changes systems, processes and services as a result of reform when interrelated concurrent processes have not yet concluded.

2.4. In light of the above complexities and interdependencies, it would greatly assist the communications sector, to understand the sequencing and interworking of these parallel processes and the 'web' of resulting online safety obligations. We are particularly keen to understand the interaction and overlap of Phase 1 Standards , the six registered Phase 1 Codes, the Phase 2 Codes work on which has just commenced (1 July 2024), the review of the Classification Scheme, the age assurance trial, the relevant parts of a revised Privacy Act (including the *Children's Online Privacy Code*) and foreshadowed hate speech and mis/disinformation legislation, as well as their operation alongside the amended BOSE.

2.5. The OSA underpins Australia's approach to regulating online safety. Consequently, we had recommended that the statutory review of the OSA take precedence over the amendment of the BOSE (amendment now finalised) and the development of the Phase 2 Codes. The rationale for amendments to the BOSE remains unclear when a review of the effectiveness and suitability of the BOSE had not yet been undertaken, and underlying definitions may be subject to change (as part of the review of the OSA) and, consequently, affect the scope of the BOSE. Likewise, it is challenging for industry to consider how it should approach the development of the Phase 2 Codes under the Online Content Scheme given that this is also an area for examination through the review.

3. Overarching comments

3.1. The feedback in our submission should be read against the background of our concerns more broadly with the approach that the OSA takes to the regulation of material online.

3.2. The objects of the OSA are to improve and promote online safety for Australians. The overarching objects of the OSA are sound and do not require alteration.

3.3. Notwithstanding this, the review ought to consider a number of factors to ensure the OSA and its review are of maximum benefit to end-users and online sector, now and in the longer term, and within a broader Australian and international context:

- Efficacy: how will the effectiveness and efficiency of Australia's online safety regulatory framework more broadly be measured, to inform which reforms may be necessary;
- Benchmarking: how can the existing regulatory obligations best be benchmarked across the online industry, to identify where regulation can deliver the greatest improvements:
- Flexibility: as technology, its use and societal expectations change, how to ensure that the regulatory framework, including the OSA, remain adaptive and robust;

- Complementarity: given the significant volume of online safety and online safety-adjacent reforms in Australia, how will the proposed OSA reforms work within the broader online and telecommunications regulatory framework, especially privacy laws in relation to age verification/assurance proposals; and
- Global implications: how does the setting of industry standards through regulation align with international legal frameworks, including, but not limited to, online safety frameworks.

3.4. Focusing on the OSA, to ensure the OSA is sustainable, fit-for-purpose, and effective, all compliance obligations should be:

- Targeted and proportionate to the level of risk emanating from the service, and the nature and severity of the potential harm;
- clear and certain in their application;
- consistent both internally and with other Australian laws;
- applied to different provider and service categories in a manner that is reasonable, evidence-based and effective, and takes account of differences in service characteristics (including distinctions between services that are public-facing and include risks involving virality of content, versus those that are not and that are subject to countervailing policy objectives such as addressing privacy, cyber-security and risk of data breach); and
- structured in a way that does not delegate critical policy decisions that should be subject to parliamentary debate and scrutiny, to a regulator without appropriate oversight or statutory guardrails.

Delegation to subordinate instruments and interpretation by the regulator

3.5. The current OSA permits – or rather facilitates – the shifting of critical decisions on scope of services, material subject to removal/blocking and, therefore, ultimately freedom of expression, to subordinate legislation that is either not subject to parliamentary scrutiny or, at best, is disallowable by Parliament – with interpretation and subsequent enforcement of requirements through an unelected office.

3.6. In fact, on various key issues the de-facto rule making power by eSafety by means of interpretation and guidance, that takes on a mandatory character, means that policy decisions that ought to be subject to parliamentary debate are not even delegated to subordinate instruments but instead remain entirely within the interpretational remit of the regulator. For instance, expectation 6 of the BOSE is very broad and requires providers to take reasonable steps to ensure that their services can be *used "in a safe manner*" and to proactively minimise "*material or activity*" that is "*unlawful or harmful*". Concepts such as 'harmful' and 'safety' are undefined. This lack of definition, in practice, results in eSafety imposing its views on scope and how the requirements ought to be satisfied. This gives rise to significant uncertainty for providers as to what the scope of the legislation actually is.

3.7. The issue is exacerbated by the fact that Australia does not have a human rights framework that could be applied to the legislation, subordinate instruments or regulator interpretation.

3.8. While the delegation of detailed rules to subordinate legislation may be acceptable for technical regulation and some consumer protections, it is inappropriate where human rights, in particular freedom of expression or invasion of privacy, are concerned. As a matter of principle, such legislation ought to be subject to full parliamentary scrutiny and debate. A human rights impact analysis as part of a Regulation Impact Statement does not adequately substitute for such scrutiny and debate.

3.9. It is, in our view, in all cases inappropriate to create undue definitional uncertainty in legislation and/or subordinate regulation, and to effectively delegate the interpretation of key concepts to a regulator.

3.10. Against this background, it is useful to note that other international online safety legislation, often cited as a reference point, apply a more nuanced risk focus and proportionality – these laws limit their application either in terms of the services covered or with respect to the material in scope, or both. They do this through a risk threshold and/or size and reach limitations, or a more targeted list of in-scope services (as opposed to an approach that brings all websites and apps operating in Australia within the scope of the legislation). They also do not, or to a lesser extent, delegate key rulemaking to subordinate legislation or the discretion of the regulator. Most jurisdictions are also subject to national human rights legislation.

Risk-based, proportionate approach

3.11. Unfortunately, the current OSA is not risk-based or proportionate. In general, the legislation does not apportion obligations on services based on factors such as their role and impact in the online ecosystem. The problem is largely rooted in an overly broad application of the OSA to all online services, i.e. the eight sections of the online industry, and definitions for those sections that do not include a risk threshold or otherwise limit the application of the OSA in a proportionate manner.

3.12. Attempts have been made in the Phase 1 Codes to apply some risk-based differentiation between services. In addition, a number of BOSE expectations require providers to take *"reasonable steps"* towards certain things which should operate to enable providers to apply measures that are reasonably adapted to the level of risk. However, this is not embedded more generally at a statutory level and in practice all providers in Australia (including all websites and apps made available in Australia not just those operated by large technology companies, or those that carry particular risks of virality) are required to comply.

3.13. This results, in practice, in a failure to apply a risk-based approach and disproportionate requirements on services carrying a low (or no) risk of harm. To take a couple of examples:

- A range of BOSE expectations (for example, the majority of expectations13 to 21) contain no 'reasonable steps' qualifier and simply apply equally to all services.
- Even where a 'reasonable steps' qualifier is applied, the eSafety Commissioner has issued significant Guidance on what the Commissioner expects providers to do to meet expectations which are couched in mandatory terms, despite the stated intention of the BOSE to provide providers with flexibility as to how to meet expectations. This lack of flexibility in practice imposes a significant compliance burden across a broad range of service categories with different risk levels and other characteristics.

3.14. A risk-based differentiation of services is important to avoid undue regulatory burden. For example, while the BOSE 'only' apply to social media services (SMS), DIS and RES, the expectations do not differentiate between the types of services within these very broad categories, or the risks associated with them. Therefore, any service within these three sectors of the online industry can be subject to a BOSE notice and is expected to meet the 'basic' online safety requirements. The argument that, in practice, only certain organisations are likely to receive a notice, does not mean that all SMS, DIS and RES services are expected to comply and, hence, face the regulatory burden associated with compliance. For example, a large furniture store with a website presence (with or without a blog/feedback option) (DIS) could be required to respond to a BOSE notice and meet the expectations in the same manner (or face the same penalties) that a provider of online pornographic services (also a DIS) would be

required to do. A forum that primarily enables users of a particular manufacturer's device to troubleshoot technical issues with other users (SMS or DIS, depending on interpretation) or an online game that enables users to play solitaire with one another with no ability to chat to other end-users (RES) is subject to the same expectations notwithstanding the fact that the purpose, functionality and user base of these services makes it highly improbable that these services will be used to disseminate illegal, harmful or abusive materials, or that they even have the functionality to do so.

3.15. In some instances, the lack of differentiation also leads to obligations that cannot be complied with. Please refer to our comments at sections 4.7 - 4.20 below.

3.16. A risk-based approach and corresponding differentiation with respect to compliance obligations are critical to ensure that online services are not subject to a disproportionate regulatory burden.

3.17. It is, therefore, also not appropriate to impose regulation on the mere chance that future technological developments or changes to user behaviours may increase the likelihood of harm. Such dynamic changes ought to be dealt with through an incorporation of well-designed risk thresholds that bring services into scope once (and only then) that threshold has been reached. Alternatively, Ministerial declaration powers could be used to bring specific services in scope if the above mechanism do not achieve the desired result.

3.18. For example, the providers of texts *[we will refer to short message service as 'texts' to avoid confusion with social media services (SMS)]*, MMS and similar messaging services (RES) are yet to be presented with evidence that these services cause significant harm in relation to class 1 material. Nevertheless, these services are covered by the Phase 1 Standards and the BOSE. They are also in scope for the Phase 2 Codes. (Also refer to our feedback in section 4.7 - 4.20 below.)

3.19. Following from the above, we recommend the OSA, BOSE and subordinate instruments move towards a more explicit risk-based model. For this to work, the OSA must be proportionate, flexible, evidence-based, and outcomes-focused. Proportionality can be achieved through a clear distinction between the level of risk posed by different types of service, for example, by creating differentiated BOSE expectations and requirements for SMS, RES and DIS services that are tailored according to a service's risk profile and how easily harmful content can be disseminated on a service instead of the assumption of a uniform risk profile for all services within an online section. This approach would allow services to consider online safety harms in the context of their specific services and ensure the OSA and subordinate instruments target the highest risk services while not imposing onerous obligations on lower risk services.

3.20. Using a revised risk-based approach:

- A retailer's website that allows users to provide customer reviews, but of which the primary purpose is the sale of goods to customers rather than facilitating interactions between users, would be differentiated from an SMS that offers functionality which helps users to create, view, repost and amplify their content as widely and as quickly as possible. Retail customers typically do not write reviews with the intention of virality, nor are they likely to post harmful content. Therefore, they pose an inherently low risk for the type of content that is regulated under the OSA.
- The systems, processes and/or procedures required to be implemented to comply with online safety regulation by an entertainment service, such as a streaming video on demand service (SVODS)(if in scope of the OSA, refer to our comments at sections 4.27 - 4.34 below), that only makes available professionally produced/classified content would be different to the systems, processes and/or procedures of a video sharing service that deals primarily in user-generated content. This is so even if both services have recommender systems designed to

personalise content based on an end-user's previous viewing history. Recommender systems are ubiquitous with many digital services and are not inherently risky, i.e. in some contexts, a recommender system may have a very limited online safety impact (e.g. when it is used to help enable customers navigate vast content catalogues). Imposing obligations on any service that uses a recommender system regardless of context has the potential to dilute the customer experience with no incremental safety benefit.

- A video sharing service that has low user numbers and was created for an educational purpose (e.g., a university forum) would not, logically, have the same obligations as a video sharing service with millions of active users with a primary purpose of providing general entertainment. Even though the former service may use user-to-user functionality to enhance the experience for users, the primary purpose of the service is to facilitate learning between students. Such a service may not have the functionality to make user content viral and/or has a very low likelihood of harm emanating from the service. Therefore, the way consumers use the service, the risks they might face and the safety features that are appropriate are fundamentally different for those service types. A proportionate regime must account for those differences.

3.21. The Issues Paper notes the *House of Representatives Select Committee on Social Media and Online Safety* examination of a statutory duty of care for social media platforms and other digital services, and international approaches that include such a duty. If the Government decided to proceed to give further consideration to approaches to 'flip the onus of responsibility', we believe that an organisational accountability standard on Australian services with the highest risk of harm (as opposed to a blanket approach to all services) would be more appropriate. An organisational accountability standard puts the onus on service providers to undertake risk assessments, having regard to the purpose and objects of the OSA, and – where a risk threshold has been met – take proportionate steps to implement safety-by-design measures to keep Australian end-users safe online. This would allow providers that have met the risk threshold to design their systems, processes and/or procedures having regard to the specific nature of their services, service features and functionalities, their user base, etc., while still ensuring those services are accountable to end-users and eSafety.

3.22. If an organisational accountability standard were to underpin the OSA, it would be key also to translate this concept into the subordinate codes and standards, and the BOSE.

3.23. If such an organisational accountability standard was indeed contemplated, it would also be important to ensure that the standard negates the need for prescriptive, technology-specific application and ever-increasing specificity in relation to types of material, behaviours or technologies deemed harmful. This is particularly the case if services are required to assess the effectiveness of their measures and update risk assessments on an ongoing basis. Importantly, the organisational accountability standard itself would introduce accountability for services that employ new technologies that could amplify harms by placing the onus on services to reassess risk as new technologies and potential harms emerge.

3.24. Removing prescriptiveness will assist with a technology neutral and scalable legal and regulatory framework that is better equipped to keep pace with technology, behavioural and societal changes.

4. Services / equipment in scope of the Act

4.1. The OSA applies to eight sections of the online industry. Those are:

1 Social media services (SMS);
2 Relevant electronic services (RES);

3 Designated internet services (DIS);
4 Internet search engine services;
5 App distribution services;
6 Internet carriage services (ISP);
7 Hosting services which host content in Australia; and
8 Manufacturers, suppliers and installation and maintenance providers of equipment for use by end-users in Australia in connection with an SMS, RES, DIS or ISP service. *[For ease of reading, any references to 'services' or 'service providers' in our submission include a reference to equipment and to manufacturers, suppliers and installers of such equipment.]*

4.2. The Issue Paper states that *"The definitions [of the eight online sections] are based on the primary purpose of the service, such as defining social interactions and the posting of content as 'social media services' and defining messaging between online users as 'relevant electronic services.'"* [emphasis added][3]

4.3. However, only the definition of SMS includes a 'sole or primary purpose' test (section 13 of the OSA). None of the other definitions (including the quasi-definition of equipment and manufacturers, suppliers and installers of equipment in section 134(h)) of the OSA reference a primary purpose, nor do they include any other test that would limit the definition to services with an increased risk profile of making specific material accessible or to distribute such material. Some include no definition at all, and a number of the definitions have potential for overlap between categories.

4.4. Consequently, in our view, the definitions of the eight online sections are often overly broad, unclear and/or poorly drafted, thereby leading to impractical outcomes and substantial difficulties of interpretation and understanding as to whether a particular service is subject to the OSA.

4.5. Work has been done in the Phase 1 Codes to address some of these difficulties at a codes level, but these issues ought to be addressed at an OSA level.

4.6. In particular, the definition of RES, DIS, internet carriage service and, depending on the interpretation taken on RES, also hosting service require revision. If a service is envisaged to fall within the remit of the OSA, consideration should be given to whether that service is already subject to regulation under an alternative framework and to avoiding imposing overlapping or inconsistent requirements.

Relevant electronic services

Texts and MMS

4.7. The definition of RES targets any form of communications services, irrespective of the means of delivery of those communications, its potential for distribution and legal and/or technical constraints as they relate to the control over communications and their content.

4.8. Following on from our overarching comment above, we believe that this definition ought to be refocused onto those communications services with the highest risk of access and dissemination of content in scope of the legislation. Importantly, the definition must account for the technical capabilities of service providers.

4.9. Carriers and carriage service providers (C/CSPs) are technically incapable of – and prohibited by statute to – access short messages (texts) or multimedia messages sent or received by their customers to analyse or block harmful content potentially contained in such messages (or the messages themselves). It is also not possible to limit access to these services without also limiting access to voice services (i.e. the making and

[3] p. 15 Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Statutory Review of the Online Safety Act 2021 Issues Paper*, April 2024

receiving of phone calls). Importantly, eSafety and Law Enforcement Agencies have informally confirmed that these services are not significantly contributing to the harm arising from the material under consideration. It is also unlikely that the use of such services for the access and distribution of illegal material is set to increase substantially due to the possibility of legal interception on those services. Arguably, other messaging services, including email, also do not significantly contribute to the risk of harm. In addition, end-to-end encrypted services are also constrained in their control and visibility over content.

4.10. We note that the analysis and blocking of scam messages by carriers is technically not comparable to the inspection and analysis that would be required to assess texts and MMS for harmful content. We also highlight that the former (scam analysis) is permitted by statute while the inspection of texts and MMS for other purposes is not.

4.11. All C/CSPs are subject to the Telecommunications Industry Ombudsman (TIO) scheme and the messaging services (and many other services) they provide are subject to the *Telecommunications (Consumer Complaints Handling) Industry Standard 2018*. This provides for a robust complaint scheme that consumers can avail themselves of for any of the services provided by C/CSPs. The *Telecommunications Consumer Protections (TCP) Code* sets out extensive information provision requirements in relation to all services provided by C/CSPs to consumers. All C/CSPs are also subject to extensive investigatory powers of the Australian Communications and Media Authority (ACMA).

4.12. It is, therefore, disproportionate to maintain an overly broad definition merely to ensure requirements in relation to complaint mechanisms, information provision, record keeping and/or reporting could be applied to the providers of such services in the OSA or subordinate regulation.

4.13. Consequently, we believe that, at a minimum, the definition ought to exclude short message services and multimedia message services. We highlight that the Issues Paper itself presumes an intent to only include online messaging in the RES definition by stating that "*defining messaging between online users as 'relevant electronic services.'*"[4] [emphasis added] However, we note that we do not argue that all online messaging services ought to be included in the RES definition.

4.14. Importantly, should the risk profile for C/CSP-provided texts and MMS in the actual accessing and dissemination of harmful content change in the future to warrant inclusion of those services into the definition of RES, the Minister could make use of the current section 13A(g) to specify such services in legislative rules.

4.15. This approach would avoid unnecessary regulatory impost on those services providers – an often-declared aim in Government's regulatory reform agendas.

4.16. We note that the current inclusion of texts and MMS in the definition of RES has led to requirements that are infeasible in their application to those services, e.g. taking reasonable steps to ensure the removal of the material from the service (cyber-bullying material targeted at an Australian child, non-consensual sharing of images, cyber-abuse material targeted at an Australian adult, online content scheme) as these services cannot remove individual pieces of content or block individual messages. (Again, the problem of infeasibility of requirements is not unique to texts and MMS but also applies to other services captured within the RES definition.)

4.17. Similarly, some expectations of the BOSE, which also apply to texts and MMS, cannot be complied with. For example, section 6(5) of the recently amended BOSE require RES to "*take reasonable steps to make available controls that give end-users the choice and autonomy to support safe online interactions*" and lists a number of examples for

[4] p. 15 Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Statutory Review of the Online Safety Act 2021 Issues Paper*, April 2024

such reasonable steps, such as blocking and muting controls, opt-in/out of specific content types or changes to privacy settings.

4.18. C/CSPs do not provide blocking or muting of individual messages or numbers. These features are provided on a device level, usually on the basis of the operating system of the device, i.e. they are not under the control of the C/CSP. Similarly, C/CSPs have no visibility or control of the content. They also cannot disable the text/MMS capabilities of their services without also disabling (i.e. cancelling) the attendant voice service. We are not aware of anything (without a warrant for lawful interception) that the C/CSP could do to prevent messages from reaching the intended recipient (even if the C/CSPs was aware of the content of any harmful messages), other than cancelling the entire service, including voice services. However, the latter measure can be inappropriate, e.g. where a mobile services is the only means of communication for a user.

4.19. There are other expectations that are similarly infeasible for C/CSPs, but form part of the BOSE as a result of the overly broad definition of RES and/or the lack of differentiation of requirements in the BOSE. Similar issues arise in the recently registered RES standard for class 1 material.

4.20. It is, in our view, unhelpful and inappropriate to rely on a 'reasonableness test' (determined by the Commissioner) for a multitude of obligations for those services. This approach creates end-user expectations that are not achievable for these services. It also raises the issue of discretionary interpretation of 'reasonableness' by the regulator that we highlighted in section 3 above. In our view, the better approach would be to remove these services from the RES definition or, at least, to differentiate expectations within the OSA, the BOSE and subordinate instruments.

Designated internet services

4.21. The definition of DIS is unhelpfully broad, does not sufficiently delineate DIS and hosting services, and includes services that ought to be excluded but have – inadvertently, so we assume – been included in the scope of the definition.

Websites and a multitude of other services

4.22. Roughly speaking, the definition of DIS applies to any website available to end-users in Australia (unless already captured under the definitions for other online sections), even if only a single end-user has access to that website (outside an immediate circle) and irrespective to the content accessible on or via that website, as the definition does not contain a primary purpose test or risk threshold.

4.23. The same applies to any app or any service that is transmitted via an internet carriage service, for example, a content delivery network (CDN), i.e. a service that makes identical copies of content available to end-users (on servers within Australia) to provide redundancy and increase speed of access to websites (that are hosted overseas). It equally applies to Internet of Things (IoT) network and service providers that enable the transmission and analysis of data from sensors (as those sensors, in turn, are also part of the equipment definition due to their use in 'connection with' an internet carriage service).

4.24. The current definition of DIS is not meaningful, is overly broad and, as a result, unworkable as evidenced by the complex additional definitional work and/or risk frameworks that both the (rejected) Phase 1 Code and the Phase 1 Standard were forced to employ to develop a set of rules for this section that is, at least in the case of the proposed industry code, sufficiently practical and can be complied with by the DIS section.

4.25. The definition ought to be amended to clearly focus on websites, apps and other services with a reasonable likelihood of generating harm (from the accessibility and/or distribution of harmful material) emanating from those websites.

4.26. We note that the exclusions of the definition ought to also exclude internet carriage services, in and of themselves, as well as the other forms of online service that fall within the other categories covered by the OSA.

Streaming video on demand services (SVODS)

4.27. As the Issues Paper correctly points out, the current definition of DIS at section 14 of the OSA does not include an "*on demand program service*". Unfortunately, the definition (section 18) applies the *Broadcasting Services Act 1992* (BSA) definition of 'on-demand services' which limits such services to those provided by commercial, subscription and national broadcasters, thereby limiting the exemption provided in the DIS definition to catch-up TV services. Consequently, on-demand video services (including streaming services) that fall outside the broadcaster definitions of the BSA do not benefit from the exclusion of the DIS definition.

4.28. This is inappropriate, as it leads to situations in which material may be subject to the OSA if delivered by a non-exempt 'non-broadcaster' while identical material delivered by a broadcaster is exempt. To the extent they are offering professionally produced classified content, all video on demand services (including streaming services) should be treated alike, and, therefore, be out of scope of the OSA. We also refer to our feedback at sections 5.1 - 5.9 in relation to the applicable regulatory regimes for user-generated content and professionally produced content.

4.29. R18+ material provided by SVODS has been appropriately classified in accordance with the *Classification (Publications, Film and Computer Games) Act 1995* (Classification Act).

4.30. This means that SVODS will provide local classification and advisories for their titles in accordance with the Classification Act, the *National Classification Code* and the *Guidelines for the Classification of Films 2012*. This may be by applying directly to the Classification Board to have a film classified, or self-classifying, e.g. by using an approved classification tool.

4.31. SVODS typically also provide a range of controls so users can manage their own and their family's viewing experience, such as dedicated profiles for children, with age-specific restrictions as set by a parent, and restrictions of specific titles using passwords or pin codes. Typically, they also offer tools that allow parents to monitor the usage of the service on child accounts.

4.32. As a result of the inadvertent application of the DIS definition to SVODS, these services are required to comply with a number of obligations under the class 1 DIS standard in a situation where this class 1 material is already prohibited under the Classification Scheme.

4.33. SVODS would also be required to comply with the yet to be drafted codes (or standards as the case may be) dealing with class 2 material, i.e. pornographic and 'high impact' material. However, SVODS do not provide access to X18+ material and already have obligations for R18+ content under the Classification Scheme.

4.34. Consequently, the definition of DIS ought to be amended to also exclude SVODS. The revised OSA ought to ensure that other (newly revised) definitions do not again inadvertently extend to SVODS, as this would duplicate existing separate compliance obligations.

Internet carriage services

4.35. The definition of internet carriage service has in the past been interpreted to also apply to wholesale providers of internet carriage services (or internet service providers, ISPs).

4.36. In many instances (e.g. for almost all data transmitted via fixed-line broadband services, and many other services), data is being transmitted from one end-user to another end-user with the involvement of wholesale ISPs. These providers do not hold any end-user relationship in relation to the transmission of that material but merely act as a conduit for the transmission between one or more retail ISPs which in turn have a customer relationship with one or both end-users of the communication. In many instances, two or more wholesale ISPs may be involved in the transmission of the material.

4.37. ISPs – wholesale and retail – do not have visibility of or control over the material transmitted over their networks. Consequently, their role is, in our view, more limited in relation to limiting access and distribution of relevant material beyond the blocking of websites at a domain level and cooperation with law enforcement. There is no meaningful additional action – beyond any action that could be required of retail ISPs – that wholesale ISPs could take due to the lack of any relationship with end-users and inability to exercise control over the transmitted content. End-users also typically do not seek or access information from wholesale ISPs in relation to online safety (or most other issues for that matter).

4.38. Consequently, the definition of 'internet carriage service' and/or the definition of 'supply of internet carriage service to the public ought' ought to be amended to exclude wholes ISPs.

4.39. Irrespective of our comments on specific definitions, we contend that the OSA ought to be reviewed more generally, to more appropriately focus on services that substantially contribute to the risk of truly harmful content actually being accessed or distributed (as opposed to a theoretical possibility of access or distribution).

5. Materials and behaviours in scope of Act

Interaction of the Act and the National Classification Scheme

5.1. The OSA's Online Content Scheme relies on the definitions of the *National Classification Scheme* (NCS), including the Classification Act, for class 1 and class 2 material.

5.2. Despite this reliance, the two schemes take different approaches to the regulation of material: while the NCS rests on the premise that *"adults should be able to read, hear, see and play what they want",* is designed to regulate material that is legal but may be objectionable[5] and is underpinned by a number of principles, guidance and the recognition of the context within which material may appear, the OSA's stated objective of the improvement of online safety lacks such underlying principles, guidance and the express recognition of the importance of context.

5.3. The reliance of the OSA on the NCS introduces two key issues:

- The application of the NCS in the online environment, including at scale, is highly problematic; and
- The application of the NCS in the online environment results in differential treatment of and outcomes for professionally produced content. This is especially

[5] Refer to p. 8, Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Public Consultation Paper: Modernising Australia's Classification Scheme - Stage 2*, April 2024: "Reforms he Classification the is designed so that "adults should be able to read, hear, see and play what they want; children should be protected from material likely to harm or disturb them; and everyone should be protected from exposure to unsolicited material that they find offensive."

the case where the application of the NCS in the online environment effectively results in outright bans of legal material as opposed to prohibitions of a certain activity in relation to that material, e.g. the public distribution of such material.

5.4. Importantly, the interaction and overlap of the OSA and the NCS are resulting in instances of dual regulation for certain services. Many of the services captured under the OSA currently are regulated under the NCS and the OSA, both in respect of digital and physical material. For example, an online bookseller selling both physical and electronic books will have obligations under both the OSA (e.g. to takedown ebooks in response to a removal notice) and the NCS (e.g. to not sell content that would be refused classification). While dual regulation might be acceptable where the content is classified by its creators prior to being made publicly available (e.g. cinematic films) as distributors of those films can rely on the classification rulings to determine how and where those films should be made available, the same is not true for user-generated content which will always be unclassified regardless of the medium used to make it available (usually online).

5.5. For example, where a user-generated video is uploaded to an SMS, in practice, prior classification by the user generating the content or the intermediary making the video available is impossible at scale due to the vast quantities of such content, as acknowledged in the recent *Public Consultation Paper: Modernising Australia's Classification Scheme - Stage 2*.[6]

5.6. As most of the services covered under the OSA are intermediary services (i.e., they are for the most part not responsible for creating or curating the content that is available on their services), their liability should reflect their role in the content supply chain. While the OSA acknowledges the role of intermediaries in the distribution of content, the NCS does not. This leaves regulated services in a position that requires them to navigate the complex patchwork of State and Federal media and content regulation that is out of step with the digital evolution of content distribution and consumption in Australia.

5.7. It also leaves many services exposed under laws that are not fit for purpose, e.g., a service provider that complies with the OSA in all respects may still be exposed under the NCS for making classifiable content available online.

5.8. Consequently, we support the Government's current proposal for the NCS to move towards a regime for professionally produced content, formally removing user-generated content from its scope. Importantly, the two regimes ought to clearly mirror the apparent policy distinction between user-generated content and professionally produced content: while the former ought exclusively to be subject to the OSA, the latter ought exclusively to be subject to the NCS.

5.9. We also support removing dual regulation for companies subject to both the OSA and the NCS that are not responsible for the creation or curation of the content that is available on their services. A service that has discharged its obligations under the OSA by implementing appropriate systems and processes to detect and remove harmful or illegal material from its services ought not be liable under the NCS for making that same content available to the public.

<u>Pornography and children's access to age-inappropriate content</u>

5.10. As noted above, the OSA regulates class 1 material (i.e., material that would be refused classification, including some forms of illegal material) and class 2 material. Class 2 material is material that would be classified as R18+ (material unsuitable for minors) and X18+ (consensual sexually explicit activity).

[6] p. 8, Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Public Consultation Paper: Modernising Australia's Classification Scheme - Stage 2*, April 2024

5.11. The policy intent behind regulating access to class 2 content is the restriction of access to content that is inappropriate for children, particularly online pornography. We acknowledge the prioritisation the Government is placing on addressing access to pornography for children, as well as community expectations that this content is not made available to children. However, the current regulatory framework has a number of challenges that should be addressed so that appropriate mechanisms can be introduced to limit children's access to this material.

5.12. The table at Annex 1, which outlines the current definitions of pornography used across the NCS, illustrates that the current legal framework does not rely on one universally accepted definition of pornography.

5.13. To the extent the OSA intends to deal with pornography, it should do so in a way that recognises that some forms of pornography may be more harmful than others. For example, pornography in a graphic novel or comic is unlikely to have the same impact as pornography in audio-visual films. For this reason, proposals to legislate access to pornography should be proportionate and apply a risk lens, and target industry participants that are responsible for high impact pornographic material (e.g. the porn industry).

5.14. We also note compliance challenges that arise from difficulties of distinguishing the consensual sharing of intimate images on a service vis-à-vis pornographic material.

5.15. Similarly, any proposal relating to age-verification/assurance ought to strike an appropriate balance of efficiency, effectiveness and the impact/risk of harm vectors, given the potential privacy and cyber-security concerns with the collection of personal information.

5.16. Under the OSA, eSafety will continue to have the power to request content that would be classified as X18+ be removed from a service or put behind a restricted access system. This approach constitutes a more proportionate approach than a blanket approach that fails to take into account these balancing criteria.

5.17. We further recommend the Government first evaluate the mandatory age assurance trial prior to incorporating additional requirements in relation to age verification/ assurance into the OSA.

<u>Expansion to specific types of class 2 material</u>

5.18. The OSA applies broadly to any online service and equipment connecting to the internet.

5.19. The OSA also provides notice and takedown schemes for cyber-bulling material targeted at an Australia child, non-consensual sharing of intimate images, cyber-abuse material targeted at an Australian adult and material that depicts abhorrent violent conduct. In addition, the OSA sets out a notice and takedown scheme for class 1 and class 2 material, with an additional remedial scheme of restricting access to some types of class 2 material.

5.20. Importantly, the BOSE, and to a certain extent also the Online Content Scheme, are not limited to specific types of materials: the BOSE apply to 'harmful' material and activity, with ample discretion for eSafety to determine whether specific types of content are, in eSafety's view, 'harmful' and, consequently, are subject to the BOSE expectations. eSafety's formal Guidance in relation to the BOSE makes clear, that eSafety interprets the harmful material subject to the expectations (section 13) to go beyond the harms in scope for the notice and takedown schemes of the OSA. With respect to the Online Content Scheme and the codes/standards developed under the Scheme, eSafety also exerts substantial influence and discretion over the 'themes' that are to be included the respective class 1 and class 2 materials.

5.21. Providers should not, in effect, be required to reach a view and apply systems, processes and policies etc. to restrict categories of material where a decision has not been taken by Parliament that the category of material should be restricted. As we discuss below, many categories of material are contested, and a provider should not be in potential breach of the OSA (or be publicly shamed for a failure to meet a BOSE expectation) because it has not applied the category or measures in relation to that category in the manner that reflects eSafety's expectations.

5.22. The Issues Paper proposes further specification of material/behaviours/technologies (hereafter jointly referred to as material) that ought to be explicitly addressed through the OSA, such as:

- cyber-flashing;
- online hate;
- volumetric (pile-on) attacks;
- technology-facilitated abuse and technology-facilitated gender-based violence;
- online abuse of public figures and those requiring an online presence as part of their employment;
- other potential online safety harms through emerging technologies, including:
  - generative artificial intelligence;
  - immersive technologies;
  - recommender systems;
  - end-to-end encryption; and
  - changes to technology models such as decentralised platforms.

5.23. We note that some of the above materials (recommender systems, AI and, to some extent, online hate) have been expressly included into expanded expectations of the BOSE. We believe that most or all of our feedback in relation to an expansion of material in the OSA equally applies to the BOSE.

5.24. For some of the proposed new categories of materials it is unclear as to why these would not be considered as already being subject to one of the existing notice and takedown schemes. For example, we believe that the scheme for cyber-abuse targeted at adults (alongside existing defamation law) sufficiently addresses cyber-flashing and online abuse of public figures and others requiring an online presence as part of their employment, or could address those with minor amendments. Similarly, volumetric attacks, be they directed at children or adults, appear to be covered by the cyber-abuse schemes. At least some forms of hate speech are also within scope for those schemes and/or the scheme dealing with abhorrent violent conduct.

5.25. Importantly, we consider that prior to any expansion of materials to be regulated through the OSA, beyond the express categories listed in BOSE expectation 13, all of the following principles ought to be satisfied/tests be applied:

- there is clear evidence that the material is indeed causing harm;
- the harm that emanates from the material is not already covered by other legislation or, in an online context, not already reasonably covered by other schemes of the OSA, registered codes and/or standards;
- the harm that emanates from the material is best covered through the OSA (or subordinate legislation) rather than other existing or new legislation, including economy-wide legislation; and

- there is a clear policy decision taken at the appropriate level as to the scope and definition of the new harm to be brought within the OSA, balancing appropriate interests, including human rights considerations.

5.26. Other considerations also ought to be taken into account in deciding whether an expansion of the types of material in scope of the OSA is appropriate, including whether:

- the creation and/or publication of the material or the engagement in a specific conduct by the end-user ought to be subject to a criminal offence or a form of civil liability;
- the material can, with the present state of technology, readily be detected through automated systems and processes; and
- an expansion would disproportionately affect legitimate communications, including as a result of service providers' efforts to minimise compliance risk.

5.27. Applying the above principles and with the additional considerations in mind, we do not believe that any of the categories of material meet the criteria for express inclusion in the OSA, either because they are already in scope through existing schemes (or could be in scope with some amendments to those schemes), or because they extend sufficiently beyond the online world and, accordingly, ought to be dealt with at a higher, economy-wide level.

5.28. This is particularly true for online hate (also refer to our comments on hate speech, sections 5.29 - 5.36), technology-facilitated abuse and emerging technologies. Generative AI, end-to-end encryption, recommender systems etc. permeate many aspects of our lives and ought to be considered in a wider context. If regulation for online hate, technology-facilitated abuse and emerging technologies is indeed deemed appropriate, then protections from these potential harms ought to apply to across society, irrespective of where harm arises from these material/behaviours. It is inappropriate to seek to rush to siloed solutions at the expense of potential fragmentation, inconsistencies and unnecessary legal complexity – all of which risk stymying innovation and investment.

Hate speech

5.29. We note that the recently amended BOSE now include at section 6(3)(i) *"having processes for detecting and addressing hate speech which breaches a service's terms of use and, where applicable breaches a service's policies and procedures and standards of conduct […]"* as an example of possible measures to be taken (amongst others) to demonstrate that a provider has met the core expectation of *"take[ing] reasonable steps to ensure that end-users are able to use the service in a safe manner"* and the additional expectation of *"take[ing] reasonable steps to proactively minimise the extent to which material or activity on the service is unlawful or harmful"*.

5.30. At this stage and pending further foreshadowed guidance on what constitutes hate speech, it remains unclear whether a service provider that had implemented measures to comply with the core and additional expectation but had either not prohibited hate speech in its terms of use or not implemented measures for detection of breaches in relation to hate speech would be considered meeting the expectation. We again note the concern that the BOSE leave a decision over such compliance considerations in the discretion of the regulator.

5.31. Irrespective of the above, the proposed extension to online hate in the OSA itself is very concerning for several reasons.

5.32. Australia's Parliament has not presently prohibited or criminalised (or defined) hate speech. Instead, anti-hate speech measures are being proscribed in an online context

only, and presently only through subordinate legislative instruments which are not subject to parliamentary debate, with an interpretation of what constitutes hate speech being left to private entities or eSafety. Defining 'hate', balancing freedom of expression and human rights, and determining a balanced approach to the issue in Australia is complex and difficult. It is for that very reason hate speech has, so far, not been addressed through comprehensive Commonwealth legislation. The Issue Paper acknowledges as much by stating:

"However, hate speech is highly contested and context dependent, and these policies are not always enforced in line with community expectations. The Australian Human Rights Commission has also acknowledged the lengthy and challenging process for seeking redress, 'especially for self-represented complainants, given the lack of explicit coverage for religious identities, the 6-month limitation period, and difficulties and costs associated with progressing complaints to the Federal Court and Federal Circuit and Family Court of Australia if the conciliation process is unsuccessful.'[7]"[8]

5.33. In our view, this clearly points to a need for Parliament to address hate speech at an economy-wide, federal level, despite (or indeed because of) the complexities and challenges (and public scrutiny) that may come with such a process. The debate over a topic so closely related to human rights and freedom of expression ought not be 'buried' in online harms legislation or, worse, subordinate legislation or regulator interpretation and activity.

5.34. We wish to make clear that we do not underestimate the importance online harms legislation and the harm that hate speech can cause, rather we consider that it would be inappropriate either to focus on online hate to the exclusion of hate, vilification and discrimination more broadly and within all societal domains, or to insulate these important matters from public scrutiny and debate.

5.35. Importantly, the debate ought to include consideration whether the act of hate includes private communications or, as currently the case in state or territory legislation that addresses components of hate speech and the *Racial Discrimination Act 1975 Cth,* confine the act of discrimination or hate to an act done in public (noting that in all circumstances such legislation provides substantial definitional guidance on hate. We note the difficulty (impossibility?) to determine whether a communication constitutes hate speech in private communications.

5.36. The Issues Paper notes that "*Australia's Online Safety Act could be amended in a variety of ways to complement broader Government measures addressing online hate.*"[9] The reference to 'broader Government measures' reflects recent statements by the Attorney-General.[10] If the OSA indeed requires further specificity in relation to online hate, such amendments ought only to be made after such 'broader Government measures' – presumably Federal hate speech legislation – has been made.

6. Duty of care

6.1. We strongly caution against importing new regulatory mechanisms from other jurisdictions without full consideration of the impact and effectiveness of the extremely broad range of obligations that have already been (and continue to be) placed on providers in the recent past.

[7] Australian Human Rights Commission, *National Anti-Racism Framework Scoping Report*, National Anti-Racism Framework Scoping Report 2022, 151, accessed 26 April 2024.
[8] p. 46/47, Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Statutory Review of the Online Safety Act 2021 Issues Paper*, April 2024
[9] ibid
[10] Media conference, Parliament House, 13 Feb 2024, https://ministers.ag.gov.au/media-centre/transcripts/media-conference-parliament-house-13-02-2024, accessed on 14 June 2024

6.2. If an organisational accountability standard was indeed included into the OSA, it ought only to apply to those services that pose the largest threat to online safety and have a reasonable extent of control over the material. We argue that some services, such as C/CSP services, ought to be excluded from such a standard.

6.3. The OSA already recognises some differences with respect to the services' ability to remove content, i.e. the notice and takedown regimes appropriately only apply to DIS, RES and SMS, with additional removal schemes for apps and hosting services. The website blocking powers directed at ISPs are, also appropriately, only reserved for the most egregious content that does not occur at scale, i.e. abhorrent violent conduct. C/CSP services are also already covered by the Australian Consumer Law, products safety legislation, specific regulation for vulnerable customers, the *Telecommunications Consumer Protections (TCP) Code* and other legislation/regulation.

7. Best interests of the child

Interaction with other processes addressing children's rights online

7.1. We agree that services directed at children (i.e. those directly targeted at children and not merely accessed by children) warrant special protections. Members of our association already invest substantial resources in children's safety, in addition to general safety and cyber security measures,

7.2. We also welcome further regulatory and legislative work – in cooperation with all relevant stakeholders – in relation to children's rights online, including the foreshadowed changes to the Privacy Act and the development of a *Children's Online Privacy Code*.

7.3. The recently amended BOSE now also contain an additional expectation to "*take reasonable steps to ensure that the best interests of the child are a primary consideration in the design and operation of any service that is likely to be accessed by children.*" There is also an expectation for services that are likely to be accessed by children to collect children's data to enable them to report to eSafety on the number of Australian active end users, broken down by children and adult users.

7.4. In this context, we are concerned about specific children's safety measures being proposed for implementation in the OSA while an overarching review of children's rights online is still taking place and/or is being foreshadowed, such as a *Children's Online Privacy Code*. We would like to see a coordinated approach to regulation, with an overarching policy and legislative framework in place. Consequently, we believe it would be more useful to first complete or at least substantially progress the children's privacy work prior to embarking on further children-specific amendments to the OSA. This will ensure that the OSA is consistent with and harmonises well with key economy-wide legislation and regulation. In this context we highlight the recent commitment by the Attorney-General to introduce a revised Privacy Act into Parliament by August 2024, with a view to enacting legislation later in the year or by February 2025.

Relationship of best interests of the child, statutory duty of care and safety-by-design

7.5. The Issues Paper highlights the best interests of the child, a statutory duty of care and safety-by-design as potential new concepts for inclusion into the OSA.

7.6. It is unclear whether all three together would be considered for inclusion, or only one or two of those and, if so, which ones. It is also not clear whether, for example, a service provider discharging of its duty of care could somehow still be found in breach of obligations to act in the best interests of the child or safety-by-design requirements. Alternatively, would a provider that had acted in the best interests of the child therefore also have complied with its duty of care? Does following the agreed safety-

by-design principles offer a 'safe harbour' in respect of a duty of care or the best interests of the child?

7.7. We urge the Department to provide further detail with respect to these principles and how they are envisaged to interact with each other.

Reference framework for the best interests of the child

7.8. If the Department insists on pursuing the 'best interests of the child' principle as proposed, we raise the following issues:

7.9. The Issues Paper appears to aim at consistency of the new expectation with Article 3 of the United Nations Convention of the Rights of the Child (UNCRC) which states that

"In all actions concerning children, […], the best interests of the child shall be a primary consideration."

7.10. Government, including eSafety, has previously also referenced[11] the UK Age Appropriate Design Code (AAD Code) as a useful model for Australian children's digital rights.

7.11. We agree that consistency with this UNCRC principle and the UK AAD Code forms a useful baseline also for children's rights online in Australia. Importantly, the AAP Code's explanation of the best interests of the child appropriately recognises that the best interests of the child, on the basis of the UNCRC, include considerations of a variety of needs alongside safety, including freedom of expression, privacy, agency to form their own views and have them heard, access to information, a right to association, play, etc. Without proper protections and proportionality, there is a risk that the OSA could be interpreted in ways that disadvantage – or even infringe on – rights in the full suite of the UNCRC.

7.12. We strongly recommend that, if this concept proceeds for inclusion in the OSA, the OSA also clearly reference established guidance or establishes guidance consistent with these existing approaches to the best interests of the child to allow service providers to align with globally recognised and leading wholistic approaches to digital children's rights.

7.13. We agree with the understanding put forward in the Consultation Paper to the amendment of the BOSE[12] earlier this year (and the AAP Code) that an analysis of the best interests of an individual child (and subsequent design considerations) is infeasible for service providers with a large user base and child-users of different ages. Consequently, it would be appropriate for the OSA to clarify that service providers must consider the best interests of the child users on their service more generally instead of the best interests of an individual child based on the particular circumstances of that child.

Services in scope for best interests of the child considerations

7.14. We have previously commented on the broad language used in this context, i.e. the expectation to apply the most restrictive default privacy and safety settings if the service is "*targeted at, or being used by, children*" [emphasis added]. The recently amended BOSE now adopt language that is similarly problematic by referring to a children's service as a "*service or a component of a service (such as an online app or game) [that] is likely to be accessed by children*".

[11] p. 152, Attorney-General's Department, *Privacy Act Review Report 2022*, Feb 2023 as accessed at https://www.ag.gov.au/rights-and-protections/publications/privacy-act-review-report, on 9 Feb 2024
[12] p.11, Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Amending the Online Safety (Basic Online Safety Expectations) Determination 2022—Consultation Paper*, Nov 2023

7.15. Many apps, services or components thereof are likely to be accessed by children – in the vast majority of cases for a short time and without any meaningful risk to children. To make the best interests of the child a primary consideration in the design and operation of such apps and services would simply not be useful or feasible for a large number of low-risk services. This holds true for the amended BOSE and would equally apply to similar amendments to the OSA.

7.16. Consider for example research websites or even the website and applications of the Australian Bureau of Statistics (ABS) – often referenced as a source for data for homework for high school students. Alternatively, consider the multitude of innovative apps and websites that assist with specific learning needs, often provided by small businesses. It is clearly not useful to design those services and apps with the best interests of the child as a primary consideration. Given the very low risks of these services, other interests ought to be the primary consideration to design the best possible service for the targeted user group.

7.17. Similarly, children are likely to access texts and MMS, either on their own phones or on phones that parents give to their children, either permanently or temporarily. As highlighted above, we are not aware of any measures that C/CSPs could take into account in designing their services 'with the best interests of the child' in mind, noting that C/CSPs usually do not sell services to children.

8. Codes/Standards & Enforcement

8.1. Communications Alliance, along with DIGI and four other industry associations, was instrumental in the development of the registered Phase 1 Codes. Most of these industry associations have recently received section 141 notices (OSA) to develop Phase 2 Codes.

8.2. Our experience gained from the code development process highlighted a number of issues that merit further consideration:

- The interaction of the NCS and the OSA makes the development of codes very difficult. (Refer to our feedback at section sections 5.1 - 5.9.)
- The extraordinary breadth and lack of specificity within the definitions of the six online sections and the absence of a risk-based framework add significantly to these difficulties – as eSafety would now be able to attest to, having developed standards for DIS and RES. (Refer to our feedback at sections 3 and 4.)
- The code-development process would be vastly improved by providing a mechanism to make codes or standards for subsections of the online industry, i.e., for online games, RES provided by C/CSPs, services offering pornographic content etc. We recommend enshrining Ministerial powers to request codes or standards for sub-sections, subject to a 30-day consultation period with the affected sub-section.
- The development of the Phase 1 Codes has taken more than 20 months. The development of the RES and DIS Phase 1 Standards by eSafety has taken more than 14 months. Our experience gained in the telecommunications sector also indicates that the development of industry codes usually takes at least 12 months. It is, therefore, unclear as to how a timeframe of 6 months (with an expectation to produce a first draft within 13 weeks) could be considered realistic for the development of codes across eight online sections, many of which include international stakeholders. In short, it does not afford industry a genuine opportunity to produce a workable code. Consequently, we submit that the timeframe for the development of industry codes under the OSA ought to be no less than 12 months.

- The current OSA does not provide sufficient guidance as to when a second round of public consultation would be required, once a code has been amended to take into account feedback received during public comment (or otherwise) and, consequently, may substantially differ from the earlier consultation version. This ought to be remedied, and, in addition, a duration for additional public comment periods ought to be stipulated.

8.3. Any registered codes are not directly enforceable. Instead, the regulator (eSafety) is required to investigate and find breach with a code, subsequently direct the affected entity to comply with the respective code and can only avail itself of more advanced enforcement options if a direction has not been complied with.

8.4. Consequently, many members (but not all) of our association recommend that (similar to our proposal to move to direct enforceability of consumer codes registered under the *Telecommunications Act 1997* (Telco Act)) also codes made under the OSA ought to be directly enforceable and, in that respect, be put on the same footing as regulator-made standards. We note that the *Communications Legislation Amendment (Combating Misinformation and Disinformation) Bill 2023* also envisaged direct enforceability for codes made under that legislation.

9. Penalties

9.1. The Issues Paper raises the question as to whether the enforcement and investigative powers of eSafety are appropriate. It appears to suggest that the penalty regime under the OSA and its subordinate legislation is insufficient to deter non-compliance. The Issues Paper appears to come to this conclusion on the basis of comparisons with international online safety legislative regimes and, accordingly, states:

"Broadly speaking however, Australia's penalties regime has not kept pace with newer regulatory regimes, such as in Ireland, the EU, and the UK, which apply significantly higher penalties, including penalties based on a percentage of a platform's global revenue".[13]

9.2. However, with respect to enforcement powers and penalty regimes – and other matters – the Issues Paper does not appear to take into account the differences between the cited online safety regimes (including but not limited to their breadth in application) and the Australian approach, nor does it account for differences in the broader legislative settings, such as consumer and privacy law. The cited international regimes are also still in their infancy and, so we believe, it is too early to draw firm conclusions, including the conclusion that Australia is 'falling behind'. We recommend a more wholistic analysis of the respective regimes for an international comparison once these regimes had time to bed down.

9.3. eSafety extensive powers include:

- investigating complaints under the four complaint and content-based schemes under the OSA. For each scheme, eSafety can also issue removal notices requiring relevant companies to remove or take reasonable steps to remove material, or in some cases require material be placed behind a restricted access system;
- compelling certain services to produce identity and contact information about end-users (without a court order) (section 194);
- issuing BOSE notices to SMS, DIS and RES, effectively requiring transparency reporting from those services. There is a lack of transparency and accountability

[13] p. 33, Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Statutory Review of the Online Safety Act 2021 Issues Paper*, April 2024

regarding how eSafety determines which services receive a BOSE notice, and how it assesses responses;

- the six registered class 1 codes also contain obligations for companies, (with similar obligations being expected for the class 2 codes). For example, the SMS Code requires certain companies to cooperate with eSafety, produce transparency reporting (either annually or on request), refer unresolved complaints to eSafety and consult with eSafety about relevant changes to technology; and
- Lastly, the eSafety can make public findings about a service's compliance or non-compliance with the OSA (absent any investigation or reporting notice) and publish a report 'naming and shaming' companies that fail to comply. If there are multiple compliance failures over a 12-month period, eSafety can also apply to the Court for a cessation of services order (i.e. an order that the service that is the subject to the order no longer be provided).

9.4. These powers are already extensive and are coupled with civil penalty provisions that are specific to the nature of the breach. For example, failure to comply with a BOSE notice can result in a maximum civil penalty of approximately $800,000 for corporations per contravention (i.e. per notice) while failure to comply with a removal notice can result in a maximum civil penalty of approximately $800,000 for corporations for each day the content is not removed.

9.5. Moving to a maximum penalty regime for all breaches of the OSA would fail to recognise the nuance and complexity of the current scheme and assumes that all breaches should be treated equally regardless of severity. For example, a penalty of 10% of total revenue is inappropriate for a service that complies with a takedown request in 36 hours instead of 24 hours as required by the OSA. While an approach that disincentives non-compliance through a high penalty regime may be appropriate in cases of egregious non-compliance, it should be the exception, not the rule. We recommend that any increase to penalties be on a sliding scale, having regard to the specific contravention, the magnitude of the harm, the providers role in it and whether the non-compliance was wilful/reckless or not.

9.6. Given the OSA has only been in force since 2022 (and the industry codes only since March/June 2023), we think it would be beneficial to understand the effectiveness of the existing regime prior to considering new penalties. We also caution against transposing penalty schemes from jurisdictions with entirely different legal and regulatory requirements into a regime like the OSA which is more piecemeal (given the various schemes and subordinate instruments) while at the same time being broader (and uncertain) in scope in terms of services and materials than most or all overseas regimes.

9.7. Lastly, suggestions that legislated penalties are the only deterrent for companies fail to acknowledge that many companies are successful because of the customer trust they have earnt. Customer trust is hard to gain and easy to lose, and reputational damage through public 'naming and shaming' (as provided for in the OSA) can have significant financial implications for firms reliant on brand perception and public goodwill, and may serve as a stronger deterrent than any financial penalty.

9.8. Importantly, there is no ability for providers to challenge a negative BOSE report or other public commentary in advance, i.e. once the report/commentary has been published, reputational damage is highly likely, irrespective of the outcome of a potential challenge in a court or through administrative law action.

9.9. Due the uncertainty around the scope of the BOSE (as highlighted in section 3), the highly specific 'guidance' that effectively takes on the role of mandatory requirements by virtue of the discretion afforded to eSafety, and the inability to challenge reports prior to reputational damage occurring, the BOSE constitute an unsatisfactory regulatory model.

10. Other issues

Operation of sections 232 and 235 of the OSA

10.1. Section 232 of the OSA attempts to put beyond doubt that the OSA does not limit the operation of the Telco Act.

10.2. Sections 235(c) and (d) stipulate that State or Territory law has no effect to the extent it would "*subject, or would have the effect (whether direct or indirect) of subjecting, an Australian internet service provider to liability (whether criminal or civil) in respect of carrying particular online content in the case where the service provider was not aware of the nature of the online content*" and "*requires, or would have the effect (whether direct or indirect) of requiring, an internet service provider to monitor, make inquiries about, or keep records of, online content carrier by the provider.*" Similar provisions exist for hosting providers.

10.3. From discussions at the time of drafting of the current OSA, we understand that the intention of these sections was the protection of the privacy of communications enshrined in the Telco Act, i.e., to prevent requirements in the OSA or subordinate legislation that would require surveillance of communications through C/CSP which are prevented from doing so under the Telco Act (and the TIA Act).

10.4. With this in mind, we believe that section 232 also ought to put beyond doubt that the application of the *Telecommunications (Interceptions and Access) Act 1979* (TIA Act) is not limited by the OSA as the TIA Act contains separate but related provisions that aim at the protection of the privacy of communications.

10.5. We also highlight with great concern that section 235 only applies to State and Territory law and not the OSA itself, nor its subordinate instruments. This has in the past (Phase 1 Codes) led to prolonged debates – which recurred in relation to the RES standard and are likely to resurface in relation to the Phase 2 Codes – over envisaged requirements that would require the surveillance of communications by C/CSPs, in so far as technically possible.

10.6. Consequently, we request that section 235 be amended to put beyond doubt that neither State or Territory legislation or equity, nor the OSA or its subordinate instruments, nor other Federal legislation have effect to the extent that they require action or subject providers to liability in a manner that section 235 sought to exclude.

10.7. We note that the Attorney-General's Department has foreshadowed a 're-write' of the TIA Act in the near future. If changes to the general principles of the privacy of communications and the authorisation regime for surveillance are contemplated, those ought to be considered as part of that review, rather than being imposed, directly or indirectly, through the OSA and its subordinate legislation.

Limitations of liability for voluntary action

10.8. We welcome the protections from civil proceedings and the limitations of liability for damages afforded by section 221(2) of the OSA to persons in compliance with relevant removal, remedial and deletion notices.

10.9. However, the OSA fails to provide similar protections for actions voluntarily taken by a service provider, or pursuant to mandatory obligations under codes and standards or pursuant to expectations under the BOSE. Consequently, we request that the OSA also provides for an express exclusion of liability, similar to section 230(c)(2) of the U.S. Communications Decency Act which provides for an exclusion of liability for "*any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent,*

harassing, or otherwise objectionable, whether or not such material is constitutionally protected".

Annex 1 – Definitions of pornography

Medium	Definition of pornography	Classification
All	Child pornography means descriptions and depictions of sexual activity involving minors under 18 and sexualised descriptions and depictions of nudity involving minors as they 'deal with matters of sex […] in such a way that they offend against standards of morality, decency and propriety generally accepted by reasonable adults.' Similarly, publications which describe or depict in a way that is likely to cause offence to a reasonable adult, a person who is, or appears to be, a child under 18 (whether the person is engaged in sexual activity or not) 'will also be classified 'RC'	RC
Films and other Materials	Pornography in films and other materials refers to material containing gratuitous, exploitative or offensive depictions of (i) sexual activity accompanied by fetishes or practices which are offensive or abhorrent; or (ii) incest fantasies or other fantasies which are offensive or abhorrent. Note: Fetish means an object, an action or a non-sexual part of the body which gives sexual gratification.	RC
Publications	Pornography in publications refers to material containing exploitative descriptions or depictions of (i) violence in a sexual context; (ii) sexual activity accompanied by fetishes or practices which are revolting or abhorrent; or (ii) incest fantasies or other fantasies which are offensive or revolting or abhorrent. Note: Fetish means an object, an action or a non-sexual part of the body which gives sexual gratification. Violence in a sexual context, as distinct from sexual violence, refers to a relationship between the elements of violence and sex/sexualised nudity. The relationship may be established by the placement, juxtaposition, style or content of images and text, and/or by a story-line.	RC
Computer Games	Pornography in computer games refers to material that includes or contains depictions of actual sexual activity, or simulated sexual activity that are explicit and realistic; includes or contains gratuitous, exploitative or offensive depictions of (i) activity accompanied by fetishes or practices which are offensive or abhorrent; or (ii) incest fantasies or other fantasies which are offensive or abhorrent. Note: Fetish means an object, an action or a non-sexual part of the body which gives sexual gratification.	RC
Films	Films (except RC films) that: (i) contain real depictions of actual sexual activity between consenting adults in which there is no violence, sexual violence, sexualised violence, coercion, sexually assaultive language, or fetishes or depictions which purposefully demean anyone involved in that activity for the enjoyment of viewers, in a way that is likely to cause offence to a reasonable adult; and (ii) are unsuitable for a minor to see.	X 18+
Publications	Publications (except RC publications and Category 2 restricted publications) that: (i) explicitly depict nudity, or describe or impliedly depict sexual or sexually related activity between consenting adults, in a way that is likely to cause offence to a reasonable adult; or	Category 1 Restricted

	(ii) describe or express in detail violence or sexual activity between consenting adults in a way that is likely to cause offence to a reasonable adult; or (iii) are unsuitable for a minor to see or read.	
Publications	Publications (except RC publications) that: (a) explicitly depict sexual or sexually related activity between consenting adults in a way that is likely to cause offence to a reasonable adult; or (b) depict, describe or express revolting or abhorrent phenomena in a way that is likely to cause offence to a reasonable adult and are unsuitable for a minor to see or read	Category 2 Restricted
Computer Games	NA	

COMMUNICATIONS
ALLIANCE LTD

Level 25
100 Mount Street
North Sydney
NSW 2060 Australia

Correspondence
PO Box 444
Milsons Point
NSW 1565

T 61 2 9959 9111
E info@commsalliance.com.au
www.commsalliance.com.au
ABN 56 078 026 507