Schedule 1 – Social Media Services Online Safety Code (Class 1A and Class 1B Material)

1 Structure

This Code is comprised of the terms of this Schedule together with the Online Safety Code (Class 1A and Class 1B Material) Head Terms (**Head Terms**).

2 Scope

2.1 Social media services

(a) This Code applies to a provider of social media services, so far as materials on that service are provided to Australian end-users.

(b) Social media services include a wide variety of unique services from community-based services with a local user base to larger platforms with international user bases.

(c) Social media services may include social networks, public media sharing networks, discussion forums, and consumer review networks, to the extent that they satisfy the criteria of a social media service as outlined in the OSA.

3 Definitions

Unless otherwise indicated, terms used in this Code have the meanings given in the Head Terms or as otherwise set out below.

3.1 Social media service

(a) For the purposes of this Code, **social media service** means an electronic service that:

(i) Satisfies the following conditions:

(A) the sole or primary purpose of the service is to enable online social interaction between 2 or more end-users;

(B) the service allows end-users to link to, or interact with, some or all of the other end-users;

(C) the service allows end-users to post material on the service;

(D) such other conditions (if any) as are set out in the legislative rules; or

(ii) is an electronic service specified in the legislative rules;

but does not include an exempt service (as defined by clause 3.2).

Note: Online social interaction does not include (for example) online business interaction.

(b) For the purposes of this Code, **online social interaction** includes online interaction that enables end-users to share material for social purposes.

Note: Social purposes does not include (for example) business purposes.

(c) In determining whether the condition set out in clause 3.1(a)(i)(A) is satisfied, disregard any of the following purposes:

(i) the provision of advertising material on the service;

(ii) the generation of revenue from the provision of advertising material on the service.

3.2 Exempt service

For the purposes of this clause 3.2, a service is an **exempt service** if:

(a) none of the material on the service is accessible to, or delivered to, one or more end-users in Australia; or

(b) the service is specified in the legislative rules.

3.3 Australian child

Australian child means an Australian end-user under the age of 18 years.

3.4 Young Australian child

young Australian child means an Australian end-user under the age of 16 years.

~~**3.5 Terrorist act**~~

~~**terrorist act** has the same meaning as in section 100.1 of the *Criminal Code Act 1995* (Cth).~~

~~**3.6 Very large social media service**~~

~~**very large social media service** means a a provider of a Tier 1 social media service with over 8 million monthly active Australian end users.~~

4 Risk profile

4.1 General requirement for risk assessment

(a) How this Code applies to a social media service depends on the risk posed to Australian end-users that class 1A and 1B material will be accessed, distributed, or stored on that service.

(b) Subject to clause 4.3 and except where the social media provider chooses to automatically assign a Tier 1 risk profile to the social media service in accordance with with section 5.2(a)(ii) of the Head Terms, a provider of a social media service must undertake a risk assessment to assess the risk posed to Australian end-users that class 1A and 1B material will be accessed, distributed, or stored on the service ~~and must:~~in accordance with clause 5, and must determine that the risk profile of the social media service is either Tier 1, Tier 2, or Tier 3.

~~(i) determine that the risk profile of the social media service is either Tier 1, Tier 2, or Tier 3.~~Note: Subject to section 5.2(a) of the Heads Terms, a Tier 1 service is one with a higher risk to Australian end-users that class 1A and 1B material will be accessed, distributed or stored on the service whereas Tier 2 represents a moderate risk of this occurring and Tier 3 services represent the lowest risk of this occurring; ~~and~~

~~(ii)~~ develop and apply a methodology and process for the risk assessment). A provider of a social media service should use clause 5 as a guide for developing an appropriate methodology for the risk assessment.

(c) If a provider of a social media service is required to conduct a risk assessment under this Code, the provider must conduct the risk assessment as soon as is reasonably practical in accordance with section 5.2(a) of the Head Terms.

4.2 Methodology used for risk assessment and documentation

~~(a)~~ If a risk assessment is required under this Code, the provider of the relevant social media service must:

(i) be able to reasonably demonstrate that the provider's risk assessment methodology is based on reasonable criteria which must at a minimum include the functionality, purpose and scale of the social media service and any other criteria that are

reasonably relevant for the purpose of determining the risk profile of the social media service under this Code; and

(ii) document its assessment of the risk profile of the service in a manner that clearly explains:

(A) the methodology used to determine the risk profile of the social media service (including the weighting given to each risk factor); and

(B) the process by which the assessment was carried out.

4.3 Certain categories of social media services are not required to undertake a risk assessment

(a) Clause 4.3(b) sets out the categories of social media services that are deemed to have a Tier 3 risk profile under the Code. A provider of a service that meets the requirements of this clause 4.3 is not required to undertake a risk assessment but must comply with the compliance measures specified for Tier 3 social media services in the table in clause 6.

(b) A provider of a social media service is deemed to be a Tier 3 service and not required to conduct a risk assessment where it meets each of the requirements:

(i) the social media service does not provide an integrated chat or messaging service; and

(ii) the purpose of the social media service is for:

(A) social interaction within a limited end user group that has a common community interest (such as within a school, or neighbourhood or university community or a social or religious organisation or charity or sporting club or association); or

(B) social interaction within a commercial or public enterprise that is limited to employees and or customers of the enterprise for the enterprise's stated purpose; and

(iii) the social media service does not enable Australian end-users to do any of the following:

(A) create a list of other end-users with whom an individual shares a connection within the system; or

(B) view and navigate a list of other end-user's individual connections; or

(C) construct a public or semi-public profile within a bounded system created by the service.

4.4 Changes to risk profile of a social media service

If a provider of a social media service:

(a) makes a change to its service such that it would no longer have its risk deemed as Tier 3 under clause 4.3; or

(b) makes a change to its service that would result in the service falling within a higher risk tier,

it must carry out a risk assessment in accordance with clause 4.1 and 4.2 above.

5 ~~Guidance on risk~~Risk assessment~~:~~ requirements and guidance

(a) This clause 5 ~~provides guidance~~applies where a provider of a social media service is required to undertake a risk assessment under clause 4.

(b) ~~When adopting a methodology and process for identifying and assessing risks, a~~A provider of a social media service ~~should~~must take into account the following matters when undertaking a risk assessment of a service:

(i) the need to be objective in evaluating the risk of harm posed to Australian end-users should class 1A and 1B material be accessed, distributed or stored on the service;

(ii) the geographical spread of the social media services operations and the age of the user base;

(iii) a forward-looking analysis of changes to the internal and external environment in which the social media services operates and their impact on the ability of a service to meet the objectives and outcomes of the Code including changes in the functionality, purpose and scale of the social media service;

(iv) whether different methodology and/or processes should be used to assess the risk for class 1A and class 1B material;

(v) the need to ensure that responsible persons with the right level of skills, experience and expertise are involved in the risk assess;

(vi) relevant local, regional and international guidance (including guidance published by eSafety) and best practices (for example, with reference to the Digital Trust & Safety Partnership 'Safe Framework'); and

(vii) relevant international laws and regulations that address the assessment of online safety risks and harms, that seek to achieve objectives and outcomes similar to those contained in this Code.

(c) Subject to clause 4.3 a provider of a social media service must determine the risk profile of a service as being either Tier 1, Tier 2 or Tier 3. Should a risk assessment indicate that the service may be in-between risk tiers, the provider must assign a higher risk profile to that service.

~~(c)~~(d) In determining the risk profile of a service, a provider of a social media service should use the following table as a guide for developing a risk assessment methodology.

Risk Factor	Tier 3	Tier 2	Tier 1
Functionality	The social media service does not provide an integrated chat or messaging service.	The social media service provides Australian end-users with an integrated chat, or messaging service that does not allow live video interaction	The social media service provides Australian end-users with an integrated chat or messaging service that enables live video interaction
Purpose	The purpose of the social media service is for: a) social interaction within a limited end user group that has a common community interest (such as within a school, or	The purpose of the social media service is to provide a forum for social interaction on a specific topic, such as to enable users to post reviews of products and services or for a limited commercial or public purpose such	The purpose of the social media service is general social interaction, and it is not designed for social interaction in a specific context or for a specific purpose.

Risk Factor	Tier 3	Tier 2	Tier 1
	neighbourhood or university community or a social or religious organisation or charity or sporting club or association); or b) social interaction within a commercial or public enterprise that is limited to employees and or customers of the enterprise for the enterprise's stated purpose.	as the crowdfunding of commercial or charitable activities or social causes or to start an online petition for social change.	
Number of Australian end-users that are monthly active account holders	1 - 500,000	500,001 - 3 million	Over 3 million
Total number of end-users that are monthly active account holders globally	1 - 5 million	5,000,0001 - 30 million	Over 30 million
Format of materials	The social media service enables sharing of text (including software code) or audio only.	The social media service enables sharing of text, audio, images, and video but not live video streaming.	The social media service enables sharing of materials in all types of formats (including live video streaming).
Discoverability	The social media service does not enable Australian end-users to do any of the following: a) create a list of other end-users with whom an individual shares a connection within the system; or b) view and navigate to a list of other end-users' individual connections; or c) construct a public or semi-public profile within a bounded system created by the service.	The social media service enables Australian end-users to do any of the following: a) create a list of other end-users with whom an individual shares a connection within the system; or b) view and navigate to a list of other end-users' individual connections; or c) construct a public or semi-public profile within a bounded system created by the service.	The social media service enables Australian end-users to do all of the following: a) create a list of other end-users with whom an individual shares a connection within the system; and b) view and navigate to a list of other end-users' individual connection; and c) construct a public or semi-public profile within a bounded system created by the service.

Note: The above methodology can be used to assess the risk profile of a service. When applying this table, each factor should be given equal weighting. A provider of a social media service may determine its risk profile using an adjusted or alternative risk assessment methodology to that outlined in clause 5(e) using this table as a guide, provided that the provider can reasonably demonstrate to eSafety that the provider's risk assessment methodology is based on reasonable criteria which must, at a minimum, include the functionality, purpose and scale of the social media service and any other criteria that are reasonably relevant for the purpose of determining the risk profile of the social media service under this Code.

6 Approach to measures and guidance for social media services

The table below contains mandatory minimum and optional compliance measures for providers of social media services, depending on their risk profile.

The table also includes guidance on the implementation of some measures. This guidance is not intended to be binding on providers but to guide them on the way in which they may choose to implement a measure.

Compliance measures may either be mandatory or optional based on the risk Tier in which the relevant social media service falls. If a measure is specified as mandatory for only some risk Tiers, it is optional for the other risk Tiers.

Tier	Mandatory minimum compliance measures	Optional Compliance measures
All Tiers	1, 5, 13	-
Tier 1	1-8, 9, 10-16, 18-32	17
~~**Very large social media services**~~	~~9 (in addition to all requirements in Tier 1)~~	
Tier 2	1-6, 11, 12, 20, 21, 23, 24,26-30, 33	17
Tier 3	-	-

7 Compliance measures for class 1A and class 1B material

Objective 1: Industry participants will take reasonable and proactive steps to create and maintain a safe online environment for Australian end-users.	
Outcome 1: Industry participants take reasonable and proactive steps to prevent access or exposure to, distribution of, and online storage of class 1A material.	
Minimum compliance measures for all social media services	**1) Notifying appropriate entities about class 1A material on their services** If a provider of a social media service: a) identifies CSEM and/or pro-terror materials on its service; and b) forms a good faith belief that the CSEM or pro-terror material is evidence of serious and immediate threat to the life or physical health or safety of an Australian adult or child (i.e., an adult or child ordinarily resident in Australia), it must report such material to an appropriate entity within 24 hours or as soon as reasonably practicable. An **appropriate entity** means foreign or local law enforcement (including, Australian federal or state police) or organisations acting in the public interest against child sexual abuse, such as the National Centre for Missing and Exploited Children (who may then facilitate reporting to law enforcement). ***Guidance:*** *A provider should seek to make a report to an appropriate entity as soon as reasonably practicable in light of the circumstances surrounding that report~~.~~, noting that the referral of materials under this measure to appropriate authorities is time-critical. For example, in some circumstances, a provider acting in good faith, may need time to investigate the authenticity of a report~~.~~, but when a report has been authenticated, an appropriate authority should be informed without delay. A provider should ensure that such report is compliant with other applicable laws such as Privacy Law.* Note: Measure 1 is intended to supplement any existing laws requiring social media service providers to report CSEM and pro-terror materials under foreign laws e.g., to report materials to the National Centre for Missing and Exploited Children and/or under State and Territory laws e.g., that require reporting of child sexual abuse to law enforcement.
Minimum compliance measures for Tier 1 and Tier 2 social media services	**2) Systems, processes and technologies for enforcement of policies prohibiting class 1A material** A provider of a Tier 1 or Tier 2 social media service must implement systems, processes and technologies that enable the provider to take appropriate enforcement action against end-users who breach terms and conditions, community standards, and/or acceptable use policies, prohibiting class 1A material. At a minimum social media service providers must have standard operating procedures that: a) specify the role of personnel in reviewing and responding to reports of class 1A materials by Australian end-users (more detail under measure 4); b) include clear internal channels for personnel in escalating, prioritising and assessing reports of class 1A material by Australian end-users; and c) provide operational guidance to personnel in relation to steps that should be taken when the service receives reports of class 1A materials by Australian end-users, including the steps that must be taken concerning the removal of class 1A materials in accordance with measure 3. ***Guidance:*** *The systems, processes and technologies required under measure 2 should be designed to enable providers of social media services to enforce policies in a proportionate, scalable and effective manner based the scope and urgency of potential harm that is related to the reported material, the efficacy of different types of intervention on the service, the type of service, and the source of reports. Enforcement processes should be supported by operational guidance that clearly informs the personnel on the steps they need to take to assess breaches of policies prohibiting class 1A materials and the actions they should take to enforce policies, within specified time frames including rapid response requirements for*

	reports of CSEM or pro-terror materials or where the physical safety of an Australian end-user is in immediate danger. Note: Providers must implement and publish policies prohibiting class 1A materials in accordance with measure 30.
Minimum compliance measures for Tier 1 and Tier 2 social media services	**3) Enforcement measures against account holders that breach policies prohibiting class 1A materials and age restrictions concerning the use of social media services by children** A provider of a Tier 1 or Tier 2 social media service must take appropriate enforcement action against end-users who breach terms and conditions, community standards, and/or acceptable use policies prohibiting class 1A material required under measure 30 or who breach age restrictions concerning the use of the service by an Australian child. Enforcement action under this measure must be reasonably proportionate to the level of harm associated with the relevant breach. A provider of a Tier 1 or Tier 2 social media service must: a) remove instances of CSEM or Pro-terror materials that are identified to be accessible or distributed by an Australian end-user on the service within 24 hours or as soon as reasonably practicable thereafter, unless otherwise required to deal with such material by law enforcement; b) remove other instances of class 1A materials that are identified to be accessible or distributed by an Australian end-user, as soon as reasonably practicable unless otherwise required to deal with unlawful class 1A materials by law enforcement; and c) terminate an end-user's account as soon as reasonably practicable in the event the end-user is: i) distributing CSEM or pro-terror material to Australian end-users with the intention to cause harm; or ii) known to be using the account in breach of age restrictions concerning use of the service by an Australian child; or iii) has repeatedly breached terms and conditions, community standards, and/or acceptable use policies prohibiting class 1A material on the service; and d) take reasonable steps to prevent an end-user that meets the requirements in sub-measure 3) c) i), from creating a new account for use of the service. ***Guidance:*** *In implementing sub-measure 3 c) i) and ii) and iii) the provider should seek to determine if a policy breach meets the criteria to remove an account and, if so, remove the account ~~within the relevant time frames. There~~without delay, noting that there may be circumstances where potential CSEM or pro-terror materials are identified by the service but require more than 24 hours to review, for example, it may take longer for human review of the materials to determine whether the relevant policy violation has occurred. In addition, providers should be aware that where they deploy hash technologies to identify CSEM or pro-terror materials, hash lists are not infallible; therefore, material identified by hash technologies may also require longer review times, given that an incorrect decision to categorise materials can have serious consequences for Australian end-users.* *Where the account-holder is a child that has breached age restrictions, providers may consider enabling the account to be reinstated when the child is of age.* *In implementing the measure required under sub-measure 3 c) iii) a provider of a Tier 1 or Tier 2 social media service should have a clear, documented policy for removing account holders for repeat violations of terms of service that sets out the conditions that should be satisfied before an account is removed.* *The reasonable steps required in sub-measure 3 d) could include, for example, detecting the end-user's device or IP address and blocking any new accounts created from that device or IP address either indefinitely or for a period of time (depending on the severity of the policy breach) or, where the service is subject to a pay wall, preventing use of a credit card known to be associated with the end-user's account to create a new account.* *A provider of a Tier 1 or Tier 2 social media service should also consider implementing the following measures for enforcement action for less serious breaches of terms and*

	conditions, community standards, and/or acceptable use policies prohibiting class 1A and class 1B material (other than CSEM or Pro-terror materials): *i) application of a "strike" or "penalty" against the end-user account;* *ii) restricting the end-user's use of their account (e.g. preventing the end-user from being able to post material);* *iii) suspending the end-user's account for a defined period.* *A provider of a Tier 1 or Tier 2 social media service should have a clear, documented policy outlining the criterion that will be used when applying any of the above measures.*
Minimum compliance measures for Tier 1 and Tier 2 social media services	**4) Trust and safety function** A provider of a Tier 1 or Tier 2 social media service must ensure that they are resourced with reasonably adequate personnel to oversee the safety of the service. Such personnel must have clearly defined roles and responsibilities, including for the operationalisation and evaluation of the systems and processes required under this Code. ***Guidance:*** *The trust and safety function may be allocated to one or more employees or external third-party service providers. Some industry participants may rely on the risk management systems of a related entity to assist with complying with this obligation.* *The trust and safety function should regularly report to the industry participant's senior management on safety issues related to the service. The trust and safety function should be subject to an adequate level of oversight and accountability by senior management and there should be clear protocols for escalating safety issues within the organisation.*
Minimum compliance measures for all social media services	**5) Safety by design assessments** If a provider of a social media service: a) has previously done a risk assessment under this Code and implements a significant new feature that may result in the service falling within a higher risk Tier; or b) has not previously done a risk assessment under this Code (due to falling into a category of service that does not require a risk assessment) and subsequently implements a significant new feature that would take it outside that category and require the provider to undertake a risk assessment under this Code, then that provider must (re)assess its risk profile in accordance with clause 4.4 of this Code and take reasonable steps to mitigate any additional risks to Australian end-users concerning material covered by this Code that result from the new feature, subject to the limitations in section 6.1 of the Head Terms. ***Guidance:*** *When conducting a safety by design assessment under this measure, the provider of the Tier 1 or Tier 2 social media service should consider whether any of the systems, processes or procedures covered by this Code concerning class 1A materials need to be updated in light of such new product or feature.* *In implementing this measure, the provider of the social media service may, for example:* *i) use the safety by design tools published by eSafety to assess the safety risks associated with a new product or feature; and* *ii) consult additional guidance related to safety risks published by eSafety.*
Minimum compliance measures for Tier 1 and Tier 2 social media services	**6) Safety by design features and settings for class 1A materials (including for children)** A provider of a Tier 1 or Tier 2 social media service must adopt appropriate features and settings that are designed to mitigate the risks to Australian end-users related to class 1A material including by anticipating and detecting safety risks posed by such material. A provider of a Tier 1 or Tier 2 social media service must at a minimum: a) implement measures that ensure that material can only be posted to or distributed on the service by a registered account-holder;

	b) make clear in terms and conditions, community standards, and/or acceptable use policies the minimum age an Australian end-user is permitted to hold an account on the service; c) take reasonable steps to prevent an Australian child that is known to be under the minimum age permitted on the service from holding an account on the service, and to remove them from the service as set out in measure 3); and d) have settings that are designed to prevent account-holders from unwanted contact from other end-users. ***Guidance:*** *The provider should also take reasonable steps to ensure that an Australian child that is less than the minimum age is not using its service. Such steps could include:* *i) requiring a user to declare their date of birth during the account registration process;* *ii) implementing age estimation technology to determine a user's age; or* *iii) using artificial intelligence tools that help to understand someone's real age.*
Minimum compliance measures for Tier 1 social media services	**7) Additional safety by design features and settings for Tier 1 social media services that permit a young Australian child to hold an account on the service** A provider of a Tier 1 social media service that permits a young Australian child to hold an account on the service must at a minimum: a) have default settings that are designed to prevent a young Australian child from unwanted contact from unknown end-users, including settings which prevent the location of the child being shared with other accounts by default; and b) easy to use tools and functionality that can help safeguard the safety of a young Australian child using the service. ***Guidance:*** *In implementing sub-measure 7(a) a provider of a Tier 1 social media service should ensure that default settings for a young Australian child permitted to hold an account on the service ensure the child cannot share their location publicly, or with end-users they are not connected with.* *All providers of social media services that permit a young Australian child to hold an account on the service should consider whether they have the maturity, capacity and capability to adopt this measure, and if not whether they should prohibit a young Australian child from holding an account on the service.*
Minimum compliance measures for Tier 1 social media services	**8) Use of systems, processes and/or technologies by Tier 1 social media services to detect and remove known CSAM** A provider of a Tier 1 social media service must deploy systems, processes and/or technologies designed to detect, flag and/or remove from the service, instances of known CSAM for example, using hashing, machine learning, artificial intelligence, or other safety technologies. At a minimum, providers of Tier 1 social media services must ensure their services use systems, processes and/or technology that: a) automatically detect and flag known CSAM, such as hash-matching technologies (for example, PhotoDNA, CSAI Match, and equivalent technology); b) prevent end-users from distributing known CSAM (for example, by 'black-holing' known URLs for such material or blocking or removing such material or preventing users from publicly posting detected material (prior to moderation); and c) identify phrases or words commonly linked to CSAM and linked activity to enable the provider to deter and reduce the incidence of such material and linked activity. ***Guidance:*** *In implementing this measure, providers of a Tier 1 social media services should carefully consider the appropriateness of different options for their services. Providers should consider the availability of different options and the capability of the provider to use those options accurately, including the need for systems and processes that prioritise the materials detected for human review, the human resourcing required to review detected materials, and the need to provide adequate health and safety arrangements for personnel*

	undertaking such review. The rights and expectations of legitimate users of social media services are also important factors for providers to consider when considering the type of technology that is appropriate for a particular service. *In implementing sub-measure 8a) a provider should be alert to the fact that hash lists are not infallible, and an errant hash can have serious consequences for Australian end-users. A provider should therefore take care to safeguard against low quality hashes and hashes prone to collisions (e.g., compilation videos) by having a suitable confirmation and quality control process to independently confirm that the material depicted in the hash is CSAM. Where a hash is likely to lead to false results, a provider should not deploy it.*
Minimum compliance measures for ~~very large~~ Tier 1 social media services	**9) Use of systems, processes and/or technologies to detect and remove certain types of pro-terror material** A provider of a ~~very large~~Tier 1 social media service will implement systems, processes and/or other technologies designed to detect, flag and/or remove ~~instances of videos and images that depict and promote a terrorist act~~known pro-terror material from the service, for example, through the use of key word searches, hashing, machine learning, or artificial intelligence that scans for videos and images that may, depending on the context, ~~depict and promote a terrorist~~be known-pro-terror material act and/or other safety technologies or systems or processes that limit users ability to post such materials on the service. ~~Nothing in this measure should be taken as referring to any steps that would break encryption or represent client-side scanning of end-users' online activity.~~ ***Guidance:*** *In implementing this measure, providers of ~~very large~~Tier 1 social media services should carefully consider the appropriateness of systems, processes and/or technological tools for their services. These may include, but are not limited to, systems that scan for hashed materials and could, for example, include proactive moderation strategies. Providers should consider the appropriateness of different options and the capability of the provider to use those options accurately, including the need for systems and processes that prioritise the materials detected for human review, the human resourcing required to review detected materials, and the need to provide adequate health and safety arrangements for personnel undertaking such review. The rights and expectations of legitimate users of social media services are also important factors for providers to consider when considering the type of technology that is appropriate for a particular service.*
Minimum compliance measures for Tier 1 social media services	**10) ~~Ongoing investment in systems, processes, and/or technologies and personnel~~Actions to be taken by Tier 1 social media services to disrupt and deter CSAM and pro-terror materials** A provider of a Tier 1 social media service must ~~make ongoing investments~~invest in systems, processes, and/or technologies ~~(for example, using hashing, machine learning,~~that aim to disrupt and deter end-users from using the service to create, post or disseminate CSAM and pro-terror material. At a minimum, a provider of a Tier 1 social media service must take the following steps: a) implement appropriate techniques that enable the provider to identify and monitor the nature of the threat and the areas of highest foreseeable risk on its services; and b) where the provider, acting reasonably, assesses that it is appropriate, feasible, and proportionate response to any risks identified in accordance with a), deploy safety technologies such as applications that utilise artificial intelligence~~, or other safety technologies) and personnel that support the capacity of the provider~~ and deep machine learning techniques, to detect~~,~~ and ~~take enforcement action concerning known~~remove new CSAM and ~~instances of videos and images~~/or new pro-terror materials on its service; or c) take alternate appropriate actions that ~~depict~~aim to deter end-users from creating, posting or disseminating CSAM and pro-terror materials via the service. Examples of appropriate actions include: d) investment in research and ~~promote a terrorist act , proportional to the incidence~~development and/or testing of ~~such~~novel technological solutions to address CSAM and/or pro-terror material ~~on the service and the extent~~for example nudging techniques targeted at deterring end-users from engaging with such materials and/or prompting users to file reports about such material;

	e) providing financial or technical support to non-governmental organisations that have recognised expertise in tackling CSAM or pro-terror material to improve their infrastructure and/or technical capabilities. f) participating in programs operated by non-governmental organisations such as Tech against Terror that are ~~accessible to Australian end-users~~designed to improve the capability of services to detect CSAM and/or pro-terror materials g) making technological solutions available to other service without charge or on an open-source basis. ***Guidance:*** h) ~~*In implementing this measure, Tier 1 social media services should prioritise investments in technology that aim to improve the tools and technologies to detect and take enforcement action against child sexual abuse material and pro-terror materials.*~~*In implementing this measure, the relevant types of Tier 1 social media services must consider that the threat to online safety posed by new CSAM and pro-terror materials is often different to the threat posed by known materials. Newly generated material is more likely to indicate current and ongoing safety risks such as against a child being groomed and coerced into producing new abusive images. All services subject to this measure should take action to monitor the risk of this material on their services, noting that the nature of the threat including its likely prevalence will vary markedly amongst different service and that the capability of a service to implement safety technologies that can reliably detect such material will also vary. In assessing the feasibility of implementing safety technologies under sub-clause 10 b) providers must assess whether such technology can be relied upon to detect new CSAM and/or pro-terror materials reliably in a manner that addresses the nature of the threat and the areas of highest foreseeable risk on the service. If not feasible, the provider must consider alterative steps that are targeted at the specific risks concerning this material on the service, which may include investing in interventions that are targeted at deterring end-users from engaging with this material*
Outcome 2: Industry participants take reasonable and proactive steps to prevent or limit access or exposure to, and distribution of class 1B material.	
Minimum compliance measures for Tier 1 and Tier 2 social media services	**11) Systems, processes and technologies for enforcement of policies** A provider of a Tier 1 or Tier 2 social media service must implement scalable, effective, systems, processes and technologies that enable the provider to take appropriate enforcement action against end-users who are known to have breached policies concerning class 1B material. At a minimum, a provider of a Tier1 or Tier 2 social media service must have standard operating procedures that: a) include clear internal channels for personnel to escalate and prioritise reports of class 1B material by an Australian end-user; and b) provide operational guidance to personnel in relation to steps that should be taken when the provider receives reports of class 1B materials by Australian end-users, including the steps that must be taken concerning the removal of class 1B materials in accordance with measure 12. ***Guidance:*** *Systems, processes, and technologies should be designed to enable providers of social media services to enforce policies in an appropriate, scalable and effective manner based on the urgency, and scope of potential harm that is related to the reported material, the efficacy of different types of intervention that are available on the service, the type of service, and the source of reports. Enforcement processes should be documented in a manner that clearly informs personnel on the steps they need to detect breaches of policies prohibiting class 1B materials and take action to enforce policies including rapid response requirements where the physical safety of an Australian end-user is in immediate danger.* Note: Providers must implement and publish policies prohibiting class 1B materials in accordance with measure 30.
Minimum compliance measures for	**12) Enforcement actions to be taken against account holders that breach policies prohibiting class 1B materials**

Tier 1 and Tier 2 social media services	A provider of a Tier 1 or Tier 2 social media service must take enforcement action against end-users who breach terms and conditions, community standards, and/or acceptable use policies prohibiting class 1B material that is proportionate to the level of harm associated with the relevant breach. A provider of a Tier 1 or Tier 2 social media service must, as soon as reasonably practicable: a) remove items of class 1B material identified on the service from the service; and b) terminate an end-user's account in the event the end-user has repeatedly breached terms and conditions, community standards, and/or acceptable use policies prohibiting class 1B material. ***Guidance:*** *In implementing sub-measure 12 a), a provider of a Tier 1 or Tier 2 social media service may experience circumstances where materials are identified by the service but need to be carefully assessed before they are removed; for example, it may take longer for human review of the materials to determine whether the relevant policy breach has occurred.* *In implementing sub-measure 12 b), a provider of a Tier 1 or Tier 2 social media service should have a clear policy for removing account holders for repeat breaches of terms of service that sets out the conditions that should be satisfied before an account is removed.* *A provider of a Tier 1 or Tier 2 social media service should also consider implementing the following measures for less serious violations of terms and conditions, community standards, and/or acceptable use policies prohibiting class 1B material that enable action to be taken where the end-user has breached such policies:* *i) application of a "strike" or "penalty" against the end-user account for each breach;* *ii) restricting the end-user's use of their account (e.g. preventing the end-user from being able to post material); or* *iii) suspending the end-user's account for a defined period.* *A provider of a Tier 1 social media service should have a clear, documented policy outlining the criterion that will be used if applying any of the above measures.*
Minimum compliance measures for all social media services	**13) Safety by design assessments** See measure 5 above.
Minimum compliance measures for Tier 1 social media services	**14) Ongoing investment in tools and personnel by Tier 1 social media services** A provider of a Tier 1 social media service must make ongoing investments in tools and personnel that support the capacity of the provider to detect and take enforcement action under this Code concerning class 1B material, proportional to the incidence of class 1B materials on the service and the extent class 1B materials are accessible to Australian end-users.
Outcome 4: Industry participants take reasonable and proactive steps to limit hosting of class 1A and 1B material in Australia.	
	This outcome does not require additional measures for social media services (see preamble to Heads of Terms).
Outcome 5: Industry participants consult, cooperate and collaborate with other industry participants in respect of the removal, disruption and/or restriction of class 1A and class 1B material.	
Minimum compliance measures for Tier 1 social media services	**15) Forum** A provider of a Tier 1 social media service must take part in an annual forum organised or facilitated by any industry association referred to in the Head Terms to discuss and evaluate the effectiveness of measures implemented under this Code and share best practice in implementing the Code and online safety in general with other industry participants.

	Note: the industry association responsible for the organisation and facilitation of the forum will ensure that the annual forum will allow online participation.
Minimum compliance measures for Tier 1 social media services	**16) Contribution to of expert groups that tackle CSEM and pro-terror material** A provider of a Tier 1 social media service must implement procedures for collaborating with eSafety, law enforcement, non-governmental or cross industry organisations that have established systems and processes that facilitate the safe, secure and lawful sharing of information that enables providers of social media services to detect and remove CSEM and pro-terror materials. ***Guidance:*** *A provider of a Tier 1 social media service should proactively engage with local and global industry and multi-stakeholder communities, coalitions, and alliances to share information and best practices for combatting CSEM and pro-terror material (for example, through open sourcing detection and moderation technologies, and by supporting research and innovation and contributing to cross-sector online safety groups and initiatives).* *Note: Providers should be aware that there may be legislation in certain jurisdictions where they operate that prevents the sharing of certain types of information.* *Examples of the type organisations that are contemplated by this measure include:* *i) the EU Internet Forum and the Global Internet Forum to Counter Terrorism for pro-terror material (for pro-terror materials); and* *ii) and the Technology Coalition, the International Centre for Missing and Exploited Children and WePROTECT Global Alliance (for CSEM).* Note: Providers should be aware some organisations can refuse membership applications.
Optional compliance measures for Tier 1 and Tier 2 social media services	**17) Working with researchers and academics** A provider of a Tier 1 or Tier 2 social media service may provide support such as funding and/or access to data for good faith research into the prevalence, impact and appropriate responses that providers of social media services may adopt in relation to class 1A and class 1B materials and the subcategories of class 1A and class 1B materials such as CSEM, and pro terror material.
Outcome 6: Industry participants communicate and cooperate with eSafety in respect of matters relating to class 1A and class 1B material, including complaints.	
Minimum compliance measures for Tier 1 social media services	**18) Timely referral of unresolved complaints to eSafety** A provider of a Tier 1 social media service must, refer to eSafety complaints from the public concerning the providers non-compliance with this Code, where the provider is unable to resolve the complaint within a reasonable time frame. ***Guidance:*** *The time frames within which providers of a Tier 1 social media service should seek to resolve complaints of non-compliance with the code and refer issues to eSafety under this measure should be based on the scope and urgency of potential harm that is related to the complaint and the source of the complaint. For example, a complaint that an Australian end-user cannot make a report of materials under this Code may take longer to resolve, where for example it requires the provider to deal with a technical problem with a reporting tool.* *Being 'unable to resolve the complaint' is intended to refer to situations where it becomes clear to the provider that their ultimate response to a given complaint is not to the satisfaction of the complainant and the complaint cannot reasonably be progressed any further between provider and complainant.*
Minimum compliance measures for Tier 1 social media services	**19) Updates and consultation with eSafety about relevant changes to technology** A provider of a Tier 1 social media service must take reasonable steps to ensure eSafety receives updates regarding significant changes to the functionality of their services that are likely to have a material positive or negative effect on the access or exposure to, distribution of, and online storage of class 1A or class 1B materials by Australian end-

	users. A provider may choose to provide this information in an annual report to eSafety under measure 32. In implementing this measure, providers are not required to disclose information to eSafety that is confidential. ***Guidance:*** *Updates to eSafety may also be provided after any public announcement of the relevant changes through mechanisms such as eSafety Advisory Committee or through reporting to eSafety.*
Objective 2: Industry participants will empower people to manage access and exposure to class 1A and class 1B material.	
Outcome 7: Industry participants provide tools and/or information to limit access and exposure to class 1A and class 1B material.	
Minimum compliance measures for Tier 1 social media services and Tier 2 social media services that have account holders who are children	**20) Provision of information to parents and carers and young Australian children** A provider of Tier 1 social media service and a provider of a Tier 2 social media service that permits a young Australian child to be an account holder, must provide clear and easily accessible information to: a) parents and carers about how to manage the child's access and exposure to class 1A and class 1B material; and b) explain the safety tools and settings on the service in a manner that is easily understood by users of all ages permitted on the service. ***Guidance:*** *In implementing this measure, a provider of a social media service should:* *i) use simple, plain, and straightforward language;* *ii) explain the type of safety risks that children may be exposed to on the service;* *iii) explain the use of parental control tools; and* *In the case of a Tier 1 social media service, include this information in a dedicated location of the services website under measure 22.*
Minimum compliance measures for providers of Tier 1 and Tier 2 social media services	**21) Information for Australian end-users about the role and functions of eSafety, including how to make a complaint to eSafety** A provider of a Tier 1 or Tier 2 social media service must publish clear information that is accessible to Australian end-users which explains the role and functions of eSafety, including how to make a complaint to eSafety.
Minimum compliance measures for providers of Tier 1 social media services	**22) Location on the service that is dedicated to providing online safety information on Tier 1 social media services** A provider of a Tier 1 social media service must establish a location on the service that is dedicated to providing online safety information, that: a) Contains information required under measure 20, 21, 23, 24, and 25; and b) include information about how Australian end-users can contact third party services that may provide counselling and support; c) is accessible to Australian end-users. ***Guidance:*** *A provider should consider raising Australian end-users' awareness about the availability of safety information on its platform in relation to its services, through interstitial mechanisms such as account notifications, on-platform advertising campaigns or pop-up notices when material is being posted or viewed by Australian end-users. Providers should also consider contributing to off-platform campaigns targeted at the general public, Australian end-users or specific sections of the community such as teachers, parents and carers, older users or vulnerable groups. A provider may also consider contributing to an off-platform campaign*

	by providing financial assistance, advertising collateral, expert advisers, or other support services.
Outcome 8: Industry participants provide clear and effective reporting and complaints mechanisms for class 1A and class 1B material.	
Minimum compliance measures for Tier 1 and Tier 2 social media services	**23) Reporting and complaints mechanisms for class 1A and class 1B material** A provider of a Tier 1 or Tier 2 social media service must provide tools which enable Australian end-users to report, flag and/or make a complaint about class 1A and class 1B material accessible on the service. Such reporting mechanisms must: a) be easily accessible and easy to use; b) be accompanied clear instructions on how to use them, as well as an overview of the reporting process; and c) ensure that the identity of the reporter is not disclosed to the reported end-user or account holder (i.e. the individual who has been reported should not be able to see the person who reported them), without the reporter's express consent).
Minimum compliance measures for Tier 1 and Tier 2 social media services	**24) Complaints about handling of reports and/or compliance with Code** A provider of a Tier 1 or Tier 2 social media service must provide tools which enable Australian end-users to make a complaint about: a) the provider's handling of reports about class 1A or class 1B material that is accessible on the service; or b) any other aspect of the provider's compliance with this Code. Such complaints tools must: c) be easily accessible and simple to use; and d) be accompanied by plain language instructions on how to use them, as well as an overview of the complaints process.
Minimum compliance measures for providers of Tier 1 social media services	**25) On-platform reporting tools for Tier 1 social media services** A provider of a Tier 1 social media service must ensure that the reporting tools referred to in measure 23 above are available and accessible to Australian end-users on-platform. ***Guidance:*** *In implementing these measures, providers of a social media service should ensure that reporting tools are integrated within the functionality of the social media service in a manner that is visible and accessible at the point the Australian end-user accesses materials posted by other end-users.*
Outcome 9: Industry participants effectively respond to reports and complaints about class 1A and class 1B material.	
Minimum compliance measures for Tier 1 and Tier 2 social media services	**26) Appropriate steps for responding to Australian end-users regarding actions taken on reports and complaints:** A provider of a Tier 1 or Tier 2 social media service must take appropriate steps to promptly respond to Australian end-users that have made reports referred to in measure 23 or complaints referred to in measure 24. At a minimum a provider of a Tier 1 or Tier 2 social media service must ensure that an Australian end-user who makes a report or complaint is informed in a reasonably timely manner of the outcome of the report or the complaint. ***Guidance:*** *The way a provider implements this measure and the timeliness of the actions required under this measure will depend on the type of material reported, the likelihood of harm that it poses to Australian end-users, the source of the report and the risk profile of the provider of the social media service. The provider should make available information to Australian end-users about indicative timeframes for responding to reports.*

Minimum compliance measures for Tier 1 and Tier 2 social media services	**27) Policies and procedures for responding to Australian end-users' reports** A provider of a Tier 1 or Tier 2 social media service must implement and document policies and procedures which detail how it gives effect to the requirements in measure 26. ***Guidance:*** *Providers should set and monitor internal targets for response times in their policies and procedures that prioritise responses and reviews of class 1A material or class 1B material that evidences an immediate risk to the physical safety to an Australian end-user.*
Minimum compliance measures for Tier 1 and Tier 2 social media services	**28) Training for personnel responding to reports** A provider of a Tier 1 or Tier 2 social media service must ensure that personnel responding to reports are trained in the social media service's policies and procedures for dealing with reports.
Minimum compliance measures for Tier 1 and Tier 2 social media services	**29) Reviews of compliance of personnel with systems and processes** A provider of a Tier 1 or Tier 2 social media service must review the effectiveness of its reporting systems and processes to ensure reports are assessed and material removed or otherwise actioned (if necessary) within reasonably expeditious timeframes, based on the level of harm the material poses to Australian end-users. Such review must occur at least annually.
Objective 3: Industry participants will strengthen transparency of, and accountability for, class 1A and class 1B material.	
Outcome 10: Industry participants provide clear and accessible information about class 1A and class 1B material.	
Minimum compliance measures for Tier 1 and Tier 2 social media services	**30) Publication of policies prohibiting class 1A and class 1B material** A provider of a Tier 1 or Tier 2 social media service must publish clear and easily accessible terms and conditions, community standards, and/or acceptable use policies, which make clear to Australian end-users that the broad categories of class 1A and class 1B material are prohibited on the service. ***Guidance:*** *The definitions of class 1A and class 1B material are complex and may not be readily understood by Australian end-users. It may therefore not be appropriate for a provider of a social media service to use these definitions in its terms and conditions, community standards and/or acceptable use policies. Instead, a provider may communicate to Australian end-users that the broad categories of material within those definitions are prohibited on the service. For example, a provider's terms and conditions may prohibit 'child sexual exploitation' generally, rather than the full definition of CSEM under this Code. In implementing this measure, a provider of a social media service should:* *i) use simple, plain, and straightforward language;* *ii) to the extent practicable, be clear about the type of material that is prohibited; and* *iii) communicate such terms and conditions, standards and/or policies to all personnel that are directly involved in their enforcement.*
Minimum compliance measures for Tier 1 social media services	**31) Information explaining how Tier 1 social media services deal with class 1A and class 1B material** A provider of a Tier 1 social media service must publish clear and accessible information that explains the actions it takes to reduce the risk of harm to Australian end-users caused by the distribution of class 1A and class 1B material on its service.
Outcome 11: Industry participants publish annual reports about class 1A and 1B material and their compliance with this Code.	
Minimum compliance	**32) Annual reporting by providers of a Tier 1 social media service**

measures for providers of Tier 1 social media services	A provider of a Tier 1 social media service must submit a Code report which as a minimum contains the following information: a) details of any risk assessment it is required to undertake pursuant to clause 4, together with information about the risk assessment methodology adopted; b) the steps that the provider has taken to comply with the applicable minimum compliance measures; c) the volume of CSEM or pro-terror material removed by the provider of the social media service; .and d) an explanation as to why these measures are appropriate. The first Code report must be submitted by the provider of the social media service to eSafety 12 months after this Code comes into effect. The provider of the social media service must submit subsequent Code reports to eSafety annually. Note: 'appropriate' has the meaning given in the Head Terms.
Minimum compliance measures for providers of Tier 2 social media services	**33) Reporting by providers of a Tier 2 social media service** Where eSafety issues a written request to a provider of a Tier 2 social media service to submit a Code report, the provider named in such request must submit to eSafety a Code report which includes the following information: a) details of any risk assessment it is required to undertake pursuant to the Code, together with information about the risk assessment methodology adopted; b) the steps that the provider has taken to comply with their applicable minimum compliance measures; and c) an explanation as to why these measures are appropriate. A provider of a Tier 2 social media service who has received such a request from eSafety is required to submit a Code report within ~~6 months~~2 months of receiving the request, but for the first request no earlier than 12 months after this Code comes into effect. A provider of a Tier 2 social media service will not be required to submit a Code report to eSafety more than once in any 12-month period. Note: 'appropriate' has the meaning given in the Head Terms.