Online Safety (Designated Internet Services—Class 1A and Class 1B Material) Industry Standard 2024

I, Julie Inman Grant, eSafety Commissioner, determine the following industry standard.

Dated

DRAFT ONLY—NOT FOR SIGNATURE

Julie Inman Grant
eSafety Commissioner

Contents

Part 1—Preliminary 1

1 Name ... 1
2 Commencement ... 1
3 Authority ... 1
4 Object of this industry standard ... 1
5 Application of this industry standard ... 1

Part 2—Interpretation 2

6 General definitions ... 2
7 Technical feasibility ... 12

Part 3—Risk assessments and risk profiles 13

8 Requirement to carry out risk assessments and determine risk profiles of designated internet services ... 13
9 Methodology, risk factors and indicators to be used for risk assessments and risk profile determinations ... 14
10 Documenting risk assessments and risk profiles ... 15

Part 4—Online safety compliance measures 17

Division 1—Preliminary 17

11 This Part not exhaustive ... 17
12 What is appropriate action? ... 17
13 Index of requirements for designated internet services ... 17

Division 2—Minimum compliance measures—general 20

14 Terms of use ... 20
15 Notification of child sexual exploitation material and pro-terror material ... 21
16 Systems and processes for responding to breaches of terms of use or community standards: class 1A material ... 21
17 Responding to breaches of terms of use or community standards—CSEM and pro-terror material ... 23
18 Responding to breaches of terms of use or community standards—class 1B material ... 24
19 Action in response to breaches of policies relating to extreme crime and violence material and class 1B material ... 25
20 Resourcing trust and safety functions ... 25
21 Detecting and removing known CSAM ... 25
22 Detecting and removing known pro-terror material ... 27
23 Disrupting and deterring CSEM and pro-terror material ... 28
24 Development programs ... 29
25 Safety features and settings—class 1A material and class 1B material ... 31
26 Referral of unresolved complaints to the Commissioner ... 32
27 Responding and referring to the Commissioner ... 32
28 Giving information about the Commissioner to end-users in Australia ... 32
29 Mechanisms for end-users and account holders to report, and make complaints about, information on designated internet services ... 33
30 Action in response to end-user reports—Tier 1 designated internet services ... 33
31 Action in response to end-user reports – other designated internet service providers ... 34
32 Policies and terms of use to be published ... 34
33 Information on actions taken by the provider ... 35

Division 3—Reporting requirements **36**
34 Commissioner may require risk assessments and other information....36
35 Reports of technical feasibility of provisions of Division 2....36
36 Notifying new features of designated internet services....37
37 Reports on outcomes of development programs....37
38 Commissioner may require compliance reports....37
39 Extension of reporting periods....39
Part 5—Miscellaneous **40**
40 Complaint resolution arrangements....40
41 Record-keeping....40

Part 1—Preliminary

1 Name

This is the *Online Safety (Designated Internet Services—Class 1A and Class 1B Material) Industry Standard 2024*.

2 Commencement

This industry standard commences on the day that is 6 months after the later of:

(a) the day after the day on which it is registered under the Act; and

(b) the day after the day on which it is registered under the *Legislation Act 2003*.

3 Authority

This industry standard is determined under section 145 of the *Online Safety Act 2021*.

4 Object of this industry standard

The object of this industry standard is to improve online safety for Australians in respect of class 1A material and class 1B material, including by ensuring that providers of designated internet services establish and implement systems, processes and technologies to manage effectively risks that Australians will solicit, generate, distribute, get access to or be exposed to class 1A material or class 1B material through the services.

5 Application of this industry standard

(1) This industry standard applies to a designated internet service, wherever it is provided from, but only so far as it is provided to end-users in Australia.

(2) If:

(a) this industry standard applies to a designated internet service; and

(b) another industry standard, or an industry code, applies to the service; and

(c) the service's predominant functionality is more closely aligned with the other industry standard or the industry code;

this industry standard does not apply to the service.

Part 2—Interpretation

Note: A number of expressions used in this industry standard are defined in the Act, including the following:

(a) child;
(b) class 1 material;
(c) class 2 material;
(d) Classification Board;
(e) Commissioner;
(f) computer game;
(g) consent;
(h) designated internet service;
(i) material;
(j) parent:
(k) posted;
(l) publication;
(m) removed;
(n) service.

6 General definitions

Definitions

(1) In this industry standard:

acceptable use policy, for a designated internet service, means the provisions of the terms of use for the service that regulate the use of the service by end-users.

account holder, for a designated internet service, means the person who is the counterparty to the agreement for the service for the provision of the service.

Example: A designated internet service may be provided to a family, where a parent or carer has the agreement with the provider of the service. The parent or carer is the account holder. Family members (including the parent or carer who is the account holder) who use the service are end-users.

Act means the *Online Safety Act 2021*.

appropriate action: see section 12.

Australian child means a child who is in Australia.

child sexual abuse material or ***CSAM*** means material that:

(a) describes, depicts, promotes or provides instruction in child sexual abuse; or
(b) is known child sexual abuse material.

child sexual exploitation material or ***CSEM*** means material that:

(a) is or includes material that promotes, or provides instruction in, paedophile activity; or
(b) is or includes:
    (i) child sexual abuse material; or

(ii) exploitative or offensive descriptions or depictions involving a person who is, appears to be or is described as a child; or

(c) describes or depicts, in a way that is likely to cause offence to a reasonable adult, a person who is, appears to be or is described as a child (whether or not the person is engaged in sexual activity);

and, in the case of a publication, also includes material that is or includes gratuitous, exploitative or offensive descriptions or depictions of:

(d) sexualised nudity; or

(e) sexual activity involving a person who is, appears to be or is described as a child.

class 1A material means class 1 material so far as it comprises:

(a) child sexual exploitation material; or

(b) pro-terror material; or

(c) extreme crime and violence material.

Note: For the definition of ***class 1 material*** see section 106 of the Act.

class 1B material means class 1 material so far as it comprises:

(a) crime and violence material (but not extreme crime and violence material); or

(b) drug-related material.

Note: For the definition of ***class 1 material*** see section 106 of the Act.

classified means classified under the *Classification (Publications, Films and Computer Games) Act 1995*.

Note: RC is a classification.

classified DIS means a designated internet service that has the sole or predominant purpose of providing general entertainment, news, or educational content, being:

(a) films or computer games that:

(i) have been classified R18^{+} or lower; or

(ii) are exempt from classification in accordance with the *Classification (Publications, Films and Computer Games) Act 1995*; or

(b) films or computer games that have not been classified but, if classified, would likely to be classified R18^{+} or lower; or

(c) books, newspapers and magazines, whether in digital or audio form, podcasts and or digital music that, if required to be classified, would likely to be classified Category 1 or lower;

and includes a service that is taken to be a classified DIS under subsection 13(2).

compliance report means a report required by section 38.

crime and violence material, in relation to a computer game, means material that is a computer game and that, without justification:

(a) promotes, incites or instructs in matters of crime or violence, or is or includes detailed instruction in, or promotion of, matters of crime or violence; or

(b) is or includes depictions of bestiality or similar practices; or

(c) depicts, expresses or otherwise deals with matters of crime, cruelty or violence in such a way that it offends against the standards of morality, decency and propriety generally accepted by reasonable adults to the extent that it should be classified RC; or

(d) is or includes depictions of violence that:

(i) have a very high degree of impact; and

(ii) are excessively frequent, prolonged, detailed or repetitive; or

(e) is or includes depictions of cruelty or realistic violence that:

(i) have a very high degree of impact; and

(ii) are very detailed; or

(f) is or includes depictions of actual sexual violence; or

(g) is or includes depictions of implied sexual violence related to incentives or rewards.

crime and violence material, in relation to a publication, means material that is, or is included in, the publication and that, without justification:

(a) promotes, incites or instructs in matters of crime or violence, or is or includes detailed instruction in matters of crime or violence; or

(b) is or includes realistic depictions of bestiality; or

(c) depicts, expresses or otherwise deals with matters of crime, cruelty or violence in such a way that it offends against the standards of morality, decency and propriety generally accepted by reasonable adults to the extent that it should be classified RC; or

(d) is or includes gratuitous, exploitative or offensive descriptions or depictions of violence that:

(i) have a very high degree of impact; and

(ii) are excessively frequent, emphasised or detailed; or

(e) is or includes gratuitous, exploitative or offensive descriptions or depictions of cruelty or real violence that:

(i) have a very high degree of impact; and

(ii) are very detailed; or

(f) is or includes gratuitous, exploitative or offensive descriptions or depictions of sexual violence.

crime and violence material, in relation to a material that is not a computer game or a publication, means material that, without justification:

(a) promotes, incites or instructs in matters of crime or violence, or is or includes detailed instruction in matters of crime or violence; or

(b) is or includes depictions of bestiality or similar practices; or

(c) depicts, expresses or otherwise deals with matters of crime, cruelty or violence in such a way that it offends against the standards of morality,

decency and propriety generally accepted by reasonable adults to the extent that it should be classified RC; or

(d) is or includes gratuitous, exploitative or offensive depictions of violence that:
  (i) have a very high degree of impact; or
  (ii) are excessively frequent, prolonged or detailed; or

(e) is or includes gratuitous, exploitative or offensive depictions of cruelty or real violence that:
  (i) have a very high degree of impact; and
  (ii) are very detailed; or

(f) is or includes gratuitous, exploitative or offensive depictions of sexual violence.

designated internet service has the meaning given by section 14(1) of the Act.

development program means a program required by section 24.

DIS means a designated internet service.

drug means a chemical, compound, or other substance or thing, that is included in Schedule 4 of the *Customs (Prohibited Imports) Regulations 1956*.

drug-related material, in relation to a computer game, means material that, without justification:

(a) depicts the unlawful use of drugs in connection with incentives or rewards; or

(b) depicts interactive, detailed and realistic use of drugs, being unlawful use; or

(c) depicts, expresses or otherwise deals with matters of drug misuse or addiction in such a way that the material offends against the standards of morality, decency and propriety generally accepted by reasonable adults to the extent that the material should be classified RC; or

(b) is or includes detailed instruction in the unlawful use of drugs.

drug-related material, in relation to a publication, means material that, without justification:

(a) depicts, expresses or otherwise deals with matters of drug misuse or addiction in such a way that the material offends against the standards of morality, decency and propriety generally accepted by reasonable adults to the extent that the material should be classified RC; or

(b) is or includes detailed instruction in the unlawful use of drugs.

drug-related material, in relation to material that is not a computer game or a publication, means material that, without justification:

(a) depicts, expresses or otherwise deals with matters of drug misuse or addiction in such a way that the material offends against the standards of morality, decency and propriety generally accepted by reasonable adults to the extent that the material should be classified RC; or

(b) is or includes detailed instruction in the unlawful use of drugs; or

(c) is or includes material promoting the unlawful use of drugs.

end-user, of a designated internet service, means a natural person who uses the service.

Example: A designated internet service may be provided to a family, where a parent or carer has the agreement with the provider of the service. The parent or carer is the account holder. Family members (including the parent or carer who is the account holder) who use the service are end-users.

end-user managed hosting service means:

(a) a designated internet service that is primarily designed or adapted to enable end-users to store or manage material; and

(b) includes a service that is taken to be an end-user managed hosting service under subsection 13(2).

Note 1: Examples of end-user managed hosting services include online file storage services, photo storage services, and other online media hosting services, including such services that include functionality to allow end-users to post or share content.

Note 2: For purposes of this industry standard, an enterprise DIS that meets this definition will be taken to be both an enterprise DIS and an end-user managed hosting service – see subsection 13(2)(c).

Note 3: An end-user managed hosting service differs from third-party hosting services (as defined in the Hosting Services Online Safety Code (Class 1A and Class 1B Material)) which have the sole or predominant purpose of supporting the delivery of another service online and which do not directly interact with end-users.

enforcement authority means:

(a) a police force or other law enforcement authority; or

(b) an organisation (including a non-government organisation) the functions of which include receiving reports of child sexual exploitation material or pro-terror material and facilitating making those reports to law enforcement authorities.

enterprise customer means the account holder under the agreement for the provision of an enterprise DIS.

Note: The enterprise customer will often make the service available to a class of end-users, such as its staff.

enterprise DIS means a designated internet service:

(a) the account holder for which is an organisation (and not an individual); and

(b) the primary functionality of which is to enable the account holder, in accordance with the terms of use for the service, to use the service for the organisation's activities, including integrating the service into the organisation's own services that are or may be made available by the organisation to the organisation's end-users; and

(c) that is of a kind that is usually acquired by account holders for the purpose mentioned in paragraph (b);

and includes a service that is taken to be an enterprise DIS under subsection 13(2).

Note 1: An enterprise DIS excludes Third-Party Hosting Services as (as defined in the Hosting Services Online Safety Code (Class 1A and Class 1B Material) and which are dealt with by that Code).

Note 2: An enterprise DIS would, for example, include:

(a) websites designed for the ordering of commercial supplies by enterprise customers; and

(b) services which provide pre-trained artificial intelligence or machine learning models for integration into a service deployed or to be deployed by an enterprise customer.

exploitative, in relation to a description or depiction of an event, means that the description or depiction:

(a) appears intended to debase or abuse, for the enjoyment of readers or viewers, the person or entity described or depicted; and

(b) has no moral, artistic or other value.

extreme crime and violence material, in relation to a computer game, means material that is crime and violence material in relation to a computer game where, without justification, the impact of the material is extreme because:

(a) the material is more detailed; or

(b) the material is realistic rather than stylised; or

(c) the game is highly interactive; or

(d) the gameplay links incentives or rewards to high impact elements of the game; or

(e) for any other reason.

extreme crime and violence material, in relation to a publication, means material that is crime and violence material in relation to a publication where, without justification, the impact of the material is extreme because of the emphasis, tone, frequency, context and detail of the relevant elements of the publication and other factors that heighten impact.

extreme crime and violence material, in relation to material that is not a computer game or a publication, means crime and violence material where, without justification, the impact of the material is extreme because:

(a) the material is more detailed; or

(b) the material is highly interactive; or

(c) the relevant depictions in the material are realistic, prolonged or repeated; or

(d) for any other reason.

general purpose DIS means a designated internet service that:

(a) is a website or application that:

(i) primarily provides information for business, commerce, charitable, professional, health, reporting news, scientific, educational, academic research, government, public service, emergency, or counselling and support service purposes;

(ii) enables transactions related to the matters in subparagraph (i); or

(b) is a web browser; and

(c) cannot be characterised as a different category of designated internet service under this industry standard.

high impact DIS means a designated internet service that:

(a) has the sole or predominant purpose of enabling end-users to access high impact materials; and

(b) makes available high impact material that has been posted by end-users;

and includes a service that is taken to be a high impact DIS because of subsection 13(2).

Note 1: This category would, for example, include websites or applications such as pornography websites and 'gore' or 'shock sites' that contain sexually explicit and/or graphically violent end-user generated content that qualifies as high impact material.

Note 2: Under paragraph 13(2)(a), a high impact DIS may also be taken to be a high impact generative AI DIS.

high impact generative AI DIS means a designated internet service:

(a) that uses machine learning models to enable an end-user to produce material; and

(b) for which it is reasonably foreseeable that the service could be used to generate synthetic high impact material.

and includes a service that is taken to be a high impact generative AI DIS because of subsection 13(2).

Note 1: This category would, for example, include services with generative artificial intelligence functionality to produce high impact material including completely new material and new material that has been created from editing existing material (for example – deepfake child sexual exploitation material).

Note 2: See note to definition of ***machine learning model platform service*** for example of an exclusion from this category.

Note 3: A high impact generative AI DIS may also be taken to be:

(a) a high impact DIS—see paragraph 13(2)(a); or

(b) a classified DIS—see paragraph 13(2)(b).

high impact materials, in relation to a high impact DIS, are materials which are:

(a) films or computer games which have been classified $R18^+$, $X18^+$ or RC in accordance with the Classification Act, or if classified would likely be classified as $R18^+$, $X18^+$ or RC; or

(b) publications which have been classified Category 1 Restricted, Category 2 Restricted, or RC in accordance with the Classification Act, or if classified would likely be classified Category 1 Restricted, Category 2 Restricted, or RC.

high impact materials, in relation to a high impact generative AI DIS are materials which are:

(a) films or computer games which have been classified $X18^+$ or RC in accordance with the Classification Act, or if classified would likely be classified as $X18^+$ or RC; or

(b) publications which have been classified Category 2 Restricted or RC in accordance with the Classification Act, or if classified would likely be classified Category 2 Restricted or RC.

industry code has the meaning given in section 132 of the Act.

justification: see subsection (4).

known child sexual abuse material means material that:

(a) is or includes images (either still images or video images); and

(b) has been verified as child sexual abuse material by a governmental (including multi-lateral) or non-governmental organisation:

(i) the functions of which are or include combating child sexual abuse or child sexual exploitation; and

(ii) in the case of a non-governmental organisation—that is generally recognised as expert or authoritative in that context; and

(c) is recorded on a database that:

(i) is managed by an organisation of a kind described in paragraph (b); and

(ii) is made available to government agencies, enforcement authorities and providers of designated internet services for the purpose of their using technological means to detect or manage child sexual abuse material on designated internet services.

Example: An example of a database referred to in paragraph (c) is the database managed by the National Center for Missing & Exploited Children.

known pro-terror material means material that has been verified as pro-terror material.

Note 1: ***Known pro-terror material*** may include material that can be detected via hashes, text signals, searches of key words terms or URLs or behavioural signals or patterns that signal or are associated with online materials produced by terrorist entities that are on the United Nations Security Council's Consolidated List.

That List was accessible, on the registration of this industry standard, at https://www.un.org/securitycouncil/content/un-sc-consolidated-list.

Note 2: Material may, for example, be verified as a result of a decision of the Classification Board. Material may also be verified by using tools provided by independent organisations that are recognised as having expertise in counter-terrorism. Examples of these organisations include Tech against Terrorism and the Global Internet Forum to Counter Terrorism.

machine learning model platform service means a designated internet service with the predominant functionality of making available one or more machine learning models and making such models available for download.

Note: A machine learning model platform service which includes functionality to enable end-users to use a hosted model to generate synthetic high impact material is not considered a high impact generative AI DIS.

offensive: see subsection (5).

pre-assessed classified DIS means a classified DIS that meets the requirements of subsection (3).

pre-assessed tier 3 designated internet service means each of the following:

(a) a pre-assessed classified DIS; and

(b) a pre-assessed general purpose DIS.

pre-assessed general purpose DIS means a general purpose DIS that meets the requirements of subsection (3).

pro-terror material means:

(a) material that:

(i) directly or indirectly counsels, promotes, encourages or urges the doing of a terrorist act; or

(ii) directly or indirectly provides instruction on the doing of a terrorist act; or

(iii) directly praises the doing of a terrorist act in circumstances where there is a substantial risk that the praise might have the effect of a leading a person (regardless of the person's age or any mental impairment that the person might suffer) to engage in a terrorist act; or

(b) material that is known pro-terror material.

However, material accessible using a designated internet service is not pro-terror material if its availability on the service can reasonably be taken to be part of public discussion, public debate, entertainment or satire.

provide a designated internet service includes make the service available.

RC means the "Refused Classification" classification under the National Classification Code.

risk assessment means an assessment of a kind required by subsection 8(1).

risk profile, for a designated internet service, means the risk profile of the service worked out under subsection 8(8).

store: material is ***stored on a designated internet service*** if it is:

(a) in storage used for the service; or

(b) accessible through or using the service.

terms of use, for a designated internet service, means the provisions of the agreement under which the service is provided and includes anything that may reasonably be regarded as the equivalent of terms of use.

terrorist act has the meaning given by section 100.1(1) of the *Criminal Code* (no matter where the action occurs, the threat of action is made or the action, if carried out, would occur).

Tier 1 designated internet service means:

(a) a designated internet service that is determined in accordance with section 8 to have a Tier 1 risk profile;
(b) a high impact DIS; and
(c) a designated internet service that is determined in accordance with section 8(9) to have a Tier 1 risk profile.

Tier 2 designated internet service means a designated internet service that is determined in accordance with section 8 to have a Tier 2 risk profile.

Tier 3 designated internet service means:
(a) a designated internet service that is determined in accordance with section 8 to have a Tier 3 risk profile; and
(b) a pre-assessed Tier 3 designated internet service.

violence means an act of violence or an obvious threat of an act of violence.

Requirements for pre-assessment

(3) The requirements for a classified DIS or a general purpose DIS to be pre-assessed are that:
(a) in respect of posting or sharing of material—the relevant service
(i) does not enable end-users in Australia to post material to the service; or
(ii) enables end-users in Australia to post material only for the purposes of enabling such end-users to review or provide information on products, services, or physical points of interest or locations made available on the service; or
(iii) enables end-users in Australia to post or share material only for the purpose of sharing that material with other end-users for a business, informational, or government service or support purpose; and
(b) in respect of chat or messaging functionality—the relevant service:
(i) does not offer a chat or messaging function; or
(ii) offers a chat or messaging function but the chat or messaging function is limited to private messages or chats between the service and end-users in Australia for a business, informational, or government service or support purpose.

Justification

(4) For this industry standard, in determining whether material is without justification, the matters to be taken into account include:
(a) the standards of morality, decency and propriety generally accepted by reasonable adults; and
(b) the literary, artistic or educational merit (if any) of the material; and
(c) the general character of the material, including whether it is of a medical, legal or scientific character; and
(d) the persons or class of persons to or amongst whom it is published or is intended or likely to be published.

Offensive material

(5) The question whether material is offensive for the purposes of this industry standard is to be determined in accordance with the Act, including section 8 of the Act.

7 Technical feasibility

In considering whether it is or is not technically feasible for the provider of a designated internet service to take a particular action, the matters to be taken into account include:

(a) the expected financial cost to the provider of taking the action; and

(b) whether it is reasonable to expect the provider to incur that cost, having regard to the extent of the risk to the online safety of end-users in Australia of not taking the action.

Part 3—Risk assessments and risk profiles

8 Requirement to carry out risk assessments and determine risk profiles of designated internet services

Risk assessments to be carried out

(1) The provider of a designated internet service must, at the times required by and in accordance with this Part, carry out an assessment of the risk that class 1A material or class 1B material:

(a) will be generated or accessed by, or distributed by or to, end-users in Australia using the service; and

(b) will be stored on the service.

Note: See also section 34.

Timing of risk assessments

(2) If the provider of the service was providing the service before the commencement of this industry standard, the risk assessment must be carried out as soon as practicable after, but no later than 6 months after, the commencement of this industry standard.

(3) Subsection (2) does not apply if a risk assessment that met the requirements of this Part had been carried out in respect of the service within 6 months before the commencement of this industry standard.

(4) A person must not start to provide a designated internet service to an end-user in Australia unless a risk assessment of the service has been carried out in accordance with this Part within 6 months before the person started to provide the service.

(5) The provider of a designated internet service must not make a material change to the service unless:

(a) a risk assessment of the service, as proposed to be changed, has been carried out in accordance with this Part; or

(b) the change will not increase the risk of class 1A material or class 1B material being accessed or generated by, or distributed to, end-users in Australia using the service, or being stored on the service.

Certain services exempt from risk assessment requirements

(6) Subsections (1) and (4) do not apply to any of the following:

(a) a pre-assessed general purpose DIS;

(b) a pre-assessed classified DIS;

(c) an end-user managed hosting service;

(d) an enterprise DIS;

(e) a high impact DIS;

(e) a high impact generative AI DIS;

(f) a machine learning model platform service;

(g) a designated internet service that is determined under subsection (9) to be a Tier 1 designated internet service.

Note: However, subsection (1) applies to a designated internet service mentioned in this subsection if the service is materially changed.

Risk profiles of designated internet services

(7) The provider of a designated internet service that conducts a risk assessment of the service must, on completion of the assessment, determine, in accordance with subsection (8), what the risk profile of the service is.

(8) The risk profile of a designated internet service is worked out as follows:

Item	If the risk that class 1A material or class 1B material will be accessed or generated by, or distributed to, end-users in Australia using the service, or will be stored on the service, is…	the risk profile of the service is …
1	High	Tier 1
2	Moderate	Tier 2
3	Low	Tier 3

Note: Some designated internet services have a pre-assessed risk profile for purposes of this industry standard. For example, a high impact DIS is pre-assessed as having a Tier 1 risk profile, and a pre-assessed classified DIS and pre-assessed general purpose DIS is each pre-assessed as having a Tier 3 risk profile.

(9) However, the provider of a designated internet service may, at any time, without having conducted a risk assessment, determine that the risk profile of the service is Tier 1.

Note: See also section 34.

9 Methodology, risk factors and indicators to be used for risk assessments and risk profile determinations

Requirement for plan and methodology

(1) If the provider is required by this Part to carry out a risk assessment for a service, the provider must formulate in writing a plan, and a methodology, for carrying out the assessment that ensure that the risks mentioned in subsection 8(1) in relation to the service are accurately assessed.

(2) The provider must ensure that the risk assessment is carried out in accordance with the plan and methodology.

(3) The provider must ensure that a risk assessment is carried out by persons with the relevant skills, experience and expertise.

Forward-looking analyses of likely changes

(4) As part of a risk assessment carried out as required by this Part, the provider must undertake a forward-looking analysis of:

(a) likely changes to the internal and external environment in which the service operates or will operate, including likely changes in the functionality of, or the scale of, the service; and

(b) the impact of those changes on the ability of the service to meet the object of this industry standard.

Note: For the object of this industry standard see section 4.

Matters to be taken into account

(5) Without limiting subsection (1), the methodology for the conduct of a risk assessment must specify the principal matters to be taken into account in assessing relevant risks, which must include the following, so far as they are relevant to the service:

(a) the predominant functionality of the service;

(b) the manner in which material is created or contributed to in connection with the service;

(c) the functionality of the service to enable end-users in Australia to post or share material;

(d) whether the service includes chat, messaging or other communications functionality;

(e) the extent to which material posted on or distributed using the service will be available to end-users of the service in Australia;

(f) the terms of use for the service;

(g) the terms of arrangements under which the provider acquires content to be made available on the service;

(h) the ages of end-users and likely end-users of the service;

(i) the outcomes of the analysis conducted as required by subsection (4);

(j) safety by design guidance and tools published or made available by a government agency or a foreign or international body;

(k) the risk to the online safety of end-users in Australia in relation to synthetic material generated by artificial intelligence.

Note 1: Arrangements referred to in paragraph (g) may include provisions that, if complied with, will reduce the risk that class 1A material and class 1B material will be made available through the service.

Note 2: Examples of agencies mentioned in paragraph (j) are the Commissioner or the Digital Trust & Safety Partnership Safe Framework.

10 Documenting risk assessments and risk profiles

(1) As soon as practicable after determining the risk profile of a designated internet service, the provider of the service must record in writing:

(a) details of the determination; and

(b) details of the conduct of any related risk assessment;

sufficient to demonstrate that they were made or carried out in accordance with this Part.

(2) The record must include the reasons for the results of the assessment and the determination of the risk profile.

Note: See also section 34.

Part 4—Online safety compliance measures

Division 1—Preliminary

11 This Part not exhaustive

This Part does not prevent the provider of a designated internet service from taking measures, in addition to and not inconsistent with those required by this Part, to improve and promote online safety for Australians.

12 What is appropriate action?

(1) In determining whether action taken or proposed in relation to a designated internet service as required by this industry standard is appropriate, the matters to be taken into account include:

(a) the extent to which the action achieves the object of this industry standard in relation to the service; and

(b) if the action relates to a breach of applicable terms of use of a designated internet service, or community standards, in relation to class 1A material or class 1B material:

(i) the nature of the material and the extent to which the breach is inconsistent with online safety for end-users in Australia; and

(ii) the extent to which the action will or may reasonably be expected to reduce or manage the risk that the service will be used to solicit, generate, access, distribute or store class 1A material or class 1B material; and

(iii) whether the proposed action is proportionate to the level of risk to online safety for end-users in Australia from the material being accessible through the service.

Note: For the object of this industry standard see section 4.

13 Index of requirements for designated internet services

(1) The following table sets out the provisions of this Part applicable to providers of designated internet services.

Item	For this kind of designated internet service …	the applicable provisions of this Part are…
1	all designated internet services	sections 34, 40 and 41
2	Tier 1 designated internet service	(a) the provisions listed in item 1 (b) sections 14, 15, 17, 19, 20, 21, 22, 25, 26, 27, 28, 29, 30, 32, 33, 35, 36 and 37 (c) subsections 16(2) and (3) (d) subsections 18(2) and (3)

Item	For this kind of designated internet service …	the applicable provisions of this Part are…
		(e) subsection 23(2) (f) subsections 24(1) to 4(a), and 24(5) to 24(6) (g) subsections 38(2) to (6), and 38(8)
3	Tier 2 designated internet service	(a) the provisions listed in item 1 (b) sections 14, 17, 19, 20, 28, 29, 31 and 32 (c) subsections 16(2) and (3) (d) subsections 18(2) and (3) (e) subsection 25(2) (f) subsections 38(2) to 38(6)
4	Tier 3 designated internet service	the provisions listed in item 1
5	end-user managed hosting service	(a) the provisions listed in item 1 (b) sections 14, 15, 17, 19, 20, 22, 26, 27, 28, 29, 31, 32, 35, 37 and 38 (c) subsections 16(2), (4), (5) and (6) (d) subsections 18(2) and (4) (e) subsections 21(2) to (8) (f) subsection 23(2) (g) subsection 25(2) (h) subsections 38(2) to (6), (8) and (9)
6	enterprise DIS	(a) the provisions listed in item 1 (b) section 14 (c) subsections 23(2) and (4) (d) subsections 38(2) to (6)
7	high impact generative AI DIS	(a) the provisions listed in item 1 (b) sections 14, 15, 17, 19, 20, 22, 24, 26, 27, 28, 29, 31, 32, 35 and 37 (c) subsections 16(2), (4) and (5) (d) subsections 18(2) and (4) (e) subsections 21(2) to (7) (f) subsections 23(2) and (3) (g) subsection 25(2) (h) subsections 38(2) to 38(6), 38(8) and 38(9)
8	machine learning model platform service	(a) the provisions listed in item 1 (b) sections 14, 15, 17, 20, 26, 27, 28, 29, 31, 37 and 38 (c) subsection 16(2) (d) subsection 23(2) (e) subsections 24(1) to (3), 24(4)(b) and 24(7) (f) subsection 25(2) (g) subsections 38(2) to (7)

Note 1: Subsection 24(4)(b) does not apply to a Tier 1 designated internet service.

Note 2: Subsection 24(4)(a) does not apply to a machine learning model platform service.

(2) Where a designated internet service meets the definition of more than one kind of designated internet service under this industry standard, for the purposes of this industry standard:

(a) if the service meets the definition of a high impact DIS and a high impact generative AI DIS—the service is taken to be a service of each of those kinds;

(b) if the service meets the definition of a classified DIS and a high impact generative AI DIS—the service will be taken to be a service of each kind;

(c) if the service meets the definition of an enterprise DIS and an end-user managed hosting service—the service will be:

(i) if made available to enterprise customers—taken to be an enterprise DIS; and

(ii) if made available by the provider directly to end-users in Australia—taken to be an end-user managed hosting service; and

(d) if the service meets the definitions of 2 or more other designated internet services—the service will be taken to be the kind of designated internet service that is most closely aligned with the service's predominant functionality.

Note 1: For paragraphs (a) and (b), this means the provider of the service must ensure the service meets the minimum compliance measures that are applicable to each kind of service.

Note 2: For paragraph (c), this means that the provider of the service must ensure the services meets the minimum compliance measures applicable to:

(a) an enterprise DIS (when the service is being provided to enterprise customers); and

(b) an end-user managed hosting service (when the service is being provided directly to end-users).

Division 2—Minimum compliance measures—general

14 Terms of use

(1) This section applies to the following services:

(a) a Tier 1 designated internet service;

(b) a Tier 2 designated internet service;

(c) an end-user managed hosting service;

(d) an enterprise DIS;

(e) a high impact generative AI DIS; and

(f) a machine learning model platform service.

Provisions to be included in terms of use

(2) The provider of a service must include provisions in the terms of use for the service that:

(a) impose an obligation on the account holder of the service to ensure that the service is not used, whether by the account holder or an end-user in Australia, to solicit, access, generate, distribute or store (as applicable, having regard to the purpose and functionality of the service) class 1A material or class 1B material; and

(b) give rights for the provider to do any of the following if the service is used to solicit, access, generate, distribute or store (as applicable) class 1A material or class 1B material:

(i) suspend the provision of the service to a specified end-user of the service for a specified period;

(ii) impose specified restrictions on the use of the service by a specified end-user of the service for a specified period;

(iii) terminate the agreement for the provision of the service.

Enforcement of terms of use

(3) If the provider of a service becomes aware of a breach of the obligation mentioned in subsection (2)(a), the provider must enforce its contractual rights in respect of the breach in an appropriate way that is proportionate to the extent of the harm to the online safety of Australians that may reasonably be expected to flow from the breach.

(4) In proceedings in respect of a contravention of subsection (3), the provider bears the evidential burden of establishing:

(a) the action it took to enforce the rights; and

(b) that the action that it took was appropriate and proportionate, as referred to in subsection (2).

Note: For appropriate action see also section 12.

15 Notification of child sexual exploitation material and pro-terror material

(1) This section applies to the following:
  (a) a Tier 1 designated internet service;
  (b) an end-user managed hosting service;
  (c) a high impact generative AI DIS;
  (d) a machine learning model platform service.

(2) If the provider of a service:
  (a) identifies child sexual exploitation material, or pro-terror material, on the service; and
  (b) believes in good faith that the material affords evidence of a serious and immediate threat to the life or physical safety of a person in Australia;

the provider must, as soon as practicable, report the matter to an enforcement authority, or otherwise as required by law.

(3) If the provider of a service:
  (a) identifies child sexual exploitation material on the service; and
  (b) believes in good faith that the material is not known child sexual exploitation material;

the provider must, as soon as practicable, notify an organisation of a kind referred to in paragraph (b) of the definition of known child sexual exploitation material in subsection 6(1).

(4) If the provider of a service:
  (a) identifies pro-terror material on the service; and
  (b) believes in good faith that the material is not known pro-terror material;

the provider must, as soon as practicable, notify an organisation that verifies material as pro-terror material.

Note: See the definition of ***pro-terror material*** in subsection 6(1).

(5) Subsections (2), (3) and (4) are in addition to any other applicable law.

16 Systems and processes for responding to breaches of terms of use or community standards: class 1A material

(1) This section applies to the following services:
  (a) a Tier 1 designated internet service;
  (b) a Tier 2 designated internet service;
  (c) an end-user managed hosting service;
  (d) a high impact generative AI DIS;
  (e) a machine learning model platform service.

Minimum requirements—generally

(2) The provider of a service must implement systems and processes that ensure that, if the provider becomes aware that:

(a) there is or has been a breach, of an obligation under the terms of use for the service in respect of class 1A material, including a breach of an obligation to comply with acceptable use policies; or

(b) there is or has been a breach, involving the service, of community standards in respect of class 1A material;

the provider takes appropriate action to ensure that:

(c) the breach, if it is continuing, ceases; and

(b) the risk of further such breaches is minimised.

Further minimum requirements—Tier 1 or Tier 2 designated internet services

(3) Without limiting subsection (2), the systems and processes implemented by a provider of a Tier 1 or Tier 2 designated internet service must include ones under which the provider:

(a) reviews reports by end-users of the service in Australia that class 1A materials are accessible using the service; and

(b) appropriately prioritises those reports and, if necessary, escalates them to senior management personnel of the provider for action.

Note: For paragraph (a), reports include the reports referred to in section 29.

Further minimum requirements —end-user managed hosting services and high impact generative AI DIS

(4) Without limiting subsection (2), a provider of an end-user managed hosting service or high impact generative AI DIS must establish standard operating procedures that:

(a) require the provider to engage with reports of class 1A material received from end-users to help determine whether the terms of use of the service (including acceptable use policies), or community standards, prohibiting class 1A material on the service have been breached; and

(b) enable the provider to take appropriate action to assess and respond to those breaches.

(5) A provider of an end-user managed hosting service or high impact generative AI DIS must implement the standard operating procedures established as required by subsection (4).

Further minimum requirements —end-user managed hosting services

(6) Without limiting subsections (2) to (5), the provider of an end-user managed hosting service must implement practices and procedures that are appropriate to minimise the likelihood that class 1A material is accessible by end-users of the service. This includes ensuring that terms of use for the service that prohibit the storage or hosting of class 1A material on the service are in place with:

(a) for an enterprise DIS—the account-holders; and

(b) in other cases—the end-users.

17 Responding to breaches of terms of use or community standards—CSEM and pro-terror material

(1) This section applies to the following:
  (a) a Tier 1 designated internet service;
  (b) a Tier 2 designated internet service;
  (c) an end-user managed hosting service;
  (d) a high impact generative AI DIS;
  (e) a machine learning model platform service.

(2) If the provider of a service becomes aware that:
  (a) there is or has been a breach of an obligation under the terms of use for the service in respect of CSEM or pro-terror material, including a breach of an obligation to comply with acceptable use policies; or
  (b) there is or has been a breach, involving the service, of community standards in respect of CSEM or pro-terror material;

  the provider must:
  (c) remove instances of CSEM and pro-terror materials identified by the provider on the service as soon as reasonably practicable unless otherwise required to deal with unlawful CSEM and pro-terror materials by an enforcement authority;
  (d) terminate an end-user's account as soon as reasonably practicable if the end-user:
    (i) is distributing CSEM or pro-terror materials to end-users with the intention to cause harm;
    (ii) is known to be an Australian child using the account; or
    (iii) has repeatedly breached terms and conditions, community standards or acceptable use policies prohibiting CSEM and pro-terror materials on the service.

(3) Without limiting subsection (2), the provider must take appropriate action to ensure that:
  (a) the service no longer permits access to or distribution of the material; and
  (b) the breach, if it is continuing, ceases; and
  (c) the risk of further such breaches is minimised; and
  (d) end-users who repeatedly breach terms of use, community standards or acceptable use policies prohibiting CSEM or pro-terror material and who have had their user accounts terminated, do not acquire new accounts.

(4) Without limiting what is appropriate action, appropriate action may include the provider exercising, in a way that is proportionate to the extent of the harm to the online safety of Australians that may reasonably be expected to flow from the breach, any of the providers contractual rights under the terms of use for the service in relation to the breach.

  Note: For the contractual rights required to be included in terms of use see paragraph 14(1)(b).

18 Responding to breaches of terms of use or community standards—class 1B material

(1) This section applies to the following:

(a) a Tier 1 designated internet service;

(b) a Tier 2 designated internet service;

(c) an end-user managed hosting service;

(d) a high impact generative AI D IS.

Minimum requirements—generally

(2) The provider of a service must implement systems and processes that ensure that, if the provider becomes aware that:

(a) there is or has been a breach, in Australia, of an obligation under for the terms of use for the service in respect of class 1B material, including a breach of an obligation to comply with acceptable use policies; or

(b) there is or has been a breach, in Australia, involving the service, of community standards in respect of class 1B material;

the provider takes appropriate action to ensure that:

(c) the breach, if it is continuing, ceases; and

(d) the risk of further such breaches is minimised.

Further minimum requirements —Tier 1 or Tier 2 designated internet services

(3) Without limiting subsection (2) the systems and processes implemented by a provider of a Tier 1 or Tier 2 designated internet service must:

(a) include ones under which the provider:

(i) reviews reports by end-users of the service in Australia that class 1B materials are accessible using the service; and

(ii) appropriately prioritises those reports and, if necessary, escalates them to senior management personnel of the provider for action; and

(b) include operational guidance to provider personnel, including actions to be take and time limits to be observed, in performing the provider's duties under this section.

Further minimum requirements for end-user managed hosting services and high impact generative AI DIS

(4) Without limiting subsection (2), the provider of an end-user managed hosting service or high impact generative AI DIS must implement standard operating procedures that:

(a) require the provider to engage with reports of class 1B material received from end-users to help determine whether the provider's terms and conditions, community standards, and/or acceptable use policies relating to class 1B materials on the service have potentially been breached and

(b) enable the provider to take appropriate action to assess and respond to potential breaches of terms and conditions, community standards, and/or acceptable use policies prohibiting class 1B material.

19 Action in response to breaches of policies relating to extreme crime and violence material and class 1B material

(1) This section applies to the following:
   (a) a Tier 1 designated internet service;
   (b) a Tier 2 designated internet service;
   (c) an end-user managed hosting service;
   (d) a high impact generative AI DIS.

(2) If the provider of a service becomes aware that:
   (a) there is or has been a breach, in Australia, of an obligation under the terms of use for the service in respect of extreme crime and violence material or class 1B material, including a breach of an obligation to comply with acceptable use policies; or
   (b) there is or has been a breach, in Australia, involving the service, of community standards in respect of extreme crime and violence material or class 1B material;

the provider must take appropriate action to respond to the breach.

20 Resourcing trust and safety functions

(1) This section applies to the following:
   (a) a Tier 1 designated internet service;
   (b) a Tier 2 designated internet service;
   (c) an end-user managed hosting service;
   (d) a high impact generative AI DIS;
   (e) a machine learning model platform service.

(2) The provider of a service must have and implement, in respect of the service, management, supervision and internal reporting arrangements to ensure that at all times the provider:
   (a) complies with the requirements of this industry standard; and
   (b) can otherwise effectively supervise the online safety of the service.

Note: These arrangements may include duties and responsibilities for personnel, and systems, processes and technologies.

(3) The provider of a service must have, or have access to, sufficient personnel who have the skills, experience and qualifications needed to ensure that the provider complies with the requirements of this industry standard at all times.

21 Detecting and removing known CSAM

(1) This section applies to the following services:

(a) a Tier 1 designated internet service;
(b) an end-user managed hosting service;
(c) a high impact generative AI DIS.

Minimum requirements—generally

(2) The provider of a service must implement systems, processes and technologies that detect and identify instances of known CSAM that:
(a) is stored on the service; or
(b) is accessible by an end-user in Australia using the service; or
(c) is being or has been accessed or distributed in Australia using the service.

Note 1: Such systems, processes and technologies include for example using hashing, machine learning, artificial intelligence systems that scan for known CSAM.

Note 2: For a high impact generative AI DIS, compliance with this subsection may require the provider to assess whether inputs into the service contain child sexual abuse material.

(3) Subsection (2) does not require a provider to implement a system, process or technology if it is not technically feasible for the provider to do so.

(4) The provider of a service must implement systems, processes and technologies that remove known CSAM from the service as soon as practicable after it is detected and identified.

(5) Subsection (4) does not require a provider to implement a system, process or technology if it is not technically feasible for the provider to do so.

(6) If it is not technically feasible for the provider to implement a particular system, process or technology for the purposes of:
(a) detecting and identifying known CSAM as required by subsection (2); or
(b) removing known CSAM as required by subsection (4);
the provider must take appropriate alternative action.

Note: For appropriate action see section 12.

(7) This section does not affect the operation of section 23.

Note 1: For technical feasibility, see section 7.

Note 2: This section does not prevent a provider from complying with legal obligations to preserve evidence of offences.

Further minimum requirements —Tier 1 designated internet services and end-user managed hosting services

(8) Without limiting subsections (2) to (7), the provider of a Tier 1 designated internet service or an end-user managed hosting service must ensure the service uses systems, processes, and technologies that automatically detect and flag known CSAM.

Further minimum requirements —Tier 1 designated internet services

(9) Without limiting subsection (2) to (8), the provider of Tier 1 designated internet service must ensure the service uses systems, processes, and technologies that:
  (a) prevent end-users from distributing known CSAM; and
  (b) 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.

22 Detecting and removing known pro-terror material

(1) This section applies to the following:
  (a) a Tier 1 designated internet service;
  (b) an end-user managed hosting service; and
  (c) a high impact generative AI DIS.

(2) The provider of a service must implement systems, processes, and technologies that detect and identify known pro-terror material that:
  (a) is stored on the service; or
  (b) is accessible by an end-user in Australia using the service; or
  (c) is being or has been generated, accessed or distributed in Australia using the service.

Note 1: Such systems, processes and technologies include for example using hashing, machine learning, artificial intelligence systems that scan for known pro-terror material.

Note 2: For a high impact generative AI DIS, compliance with this subsection may require the provider to assess whether inputs into the service contain pro-terror material.

(3) Subsection (2) does not require a provider to implement a system, process or technology if it is not technically feasible for the provider to do so.

(4) The provider of a service must implement systems, processes and technologies that remove instances of known pro-terror material from the service as soon as practicable after it is detected and identified.

(5) Subsection (4) does not require a provider to implement a system, process or technology if it is not technically feasible for the provider to do so.

(6) If it is not technically feasible for the provider to implement a particular system, process or technology for the purposes of:
  (a) detecting and identifying known pro-terror material as required by subsection (2); or
  (b) removing known pro-terror material as required by subsection (4);

the provider must take appropriate alternative action.

Note: For appropriate action see section 12.

(7) This section does not affect the operation of section 23.

Note 1: For technical feasibility, see section 7.

Note 2: This section does not prevent a provider from complying with legal obligations to preserve evidence of offences.

23 Disrupting and deterring CSEM and pro-terror material

(1) This section applies to the following:
  (a) a Tier 1 designated internet service;
  (b) an end-user managed hosting service;
  (b) an enterprise DIS;
  (c) a high impact generative AI DIS;
  (d) a machine learning model platform service.

Minimum requirements—generally

(2) The provider of a service must implement systems, processes and technologies that:
  (a) effectively deter end-users of the service from using the service; and
  (b) effectively disrupt attempts by end-users of the service to use the service;

to solicit, generate, access, distribute, store or otherwise make available CSAM and pro-terror material (including known CSAM and known pro-terror material).

Note: Examples of systems, processes and technologies include hashing, machine learning, artificial intelligence systems that scan for known CSAM and those that are designed to detect key words, behavioural signals and patterns associated with child sexual abuse material.

Further minimum requirements —high impact generative AI DIS

(3) Without limiting subsection (2), the provider of a high impact generative AI DIS must, at a minimum:
  (a) implement systems, processes and technologies that prevent generative AI features from being used to generate outputs that contain CSEM and pro-terror material;
  (b) regularly review and test models on the potential risk that a model is used to generate CSEM and pro-terror material;
  (c) promptly following review and/or testing, adjust models and deploy mitigations with the aim of reducing the misuse and unintentional use of models to generate CSEM and pro-terror material;
  (d) implement systems, processes and technologies that differentiate AI outputs generated by the model;
  (e) ensure that end-users in Australia specifically seeking images of CSAM are presented with prominent messaging that outlines the potential risk and criminality of accessing CSAM; and
  (f) ensure that material generated for end-users in Australia prompts using terms that have known associations to CSEM are accompanied by information or links to services that assist end-users in Australia to report CSEM to enforcement agencies and/or seek support; and
  (g) ensure that the systems and processes implemented by the provider under subsection (2) are able to automatically detect and action CSAM in training

data, user prompts, and outputs, with the aim of preventing this material from being generated, for example, using hashing, key word lists, classifiers, or other safety technologies.

Note: A requirement to put in place systems, processes, and technologies to disrupt and deter the production of CSEM should take account of the fact that not all high impact generative AI DIS providers will always have sufficient visibility and control of their models – if a provider lacks such visibility or control of certain aspects such that it cannot deploy all mitigations, it can rely on other systems, processes and technologies which are available.

Note: For paragraph (d), systems, processes and technologies may include by embedding indicators of provenance into material generated by a model to enable differentiation.

Further minimum requirements —enterprise DIS

(4) Without limiting subsection (2), the provider of an enterprise DIS which provides pre-trained machine learning models for integration into a service deployed or to be deployed by an enterprise customer must, at a minimum:

(a) implement and use systems, processes and technologies that automatically detect and flag CSAM in training data; and

(b) take appropriate action to ensure the service cannot be used to generate CSAM based on, using or otherwise related to such CSAM.

Note: See note 2(b) to the definition of ***enterprise DIS***.

24 Development programs

(1) This section applies:

(a) to the following:

(i) a Tier 1 designated internet service;

(ii) a high impact generative AI DIS;

(iii) a machine learning model platform service;

but only where the average monthly number of active end-users of the service, in Australia, over the immediate previous calendar year was 1,000,000 or more; and

(b) to an end-user managed hosting service but only where the average monthly number of active end-users of the service, in Australia, over the immediate previous calendar year was 500,000 or more.

(2) However:

(a) the provisions of this section, so far as they relate to pro-terror material, do not apply to a Tier 1 designated internet service predominantly used for making pornography available;

(b) paragraph 4(b) does not apply to a Tier 1 designated internet service; and

(c) paragraph 4(a) does not apply to a machine learning model platform service.

(3) The provider of the service must establish and implement, for the calendar year, a program of investment and development activities (***development program***) in respect of systems, processes and technologies.

Note: See also section 37.

(4) A development program must include:

(a) investments and activities designed to develop systems, processes and technologies that enhance the ability of the provider, or of other providers of designated internet services:

(i) to detect and identify child sexual abuse material or pro-terror material (including known child sexual abuse material and known pro-terror material) on the service; and

(ii) effectively to deter end-users of the service from using the service, and to disrupt attempts by end-users of the service to use the service, to generate, access, distribute or store child sexual abuse material or pro-terror material (including known child sexual abuse material and known pro-terror material); and

(b) arrangements for cooperating and collaborating with other organisations in activities of the kind referred to in paragraph (a) and to enhance online safety for Australians.

(5) A development program may include arrangements for the provider to make available to other providers of designated internet services, or organisations engaged in promoting online safety for Australians, systems, processes and technologies of a kind referred to in paragraph (4)(a) (including making them available without charge).

(6) Examples of investments and activities that may be part of a provider's development program for purposes of paragraph 4(a) include:

(a) procuring online safety systems and technologies for use in connection with the service, or enhancing online safety systems and technologies used in connection with the service; and

(b) conducting research into and development of online safety systems and technologies; and

(c) providing support, either financial or in kind, to organisations the functions of which are or include working to combat child sexual abuse, child sexual exploitation or terrorism.

Note: For paragraph (c), other organisations can include universities, the CSIRO, the WePROTECT Global Alliance, and the Global Internet Forum to Counter Terrorism (GIFCT).

(7) Examples of arrangements that may be part of a provider's development program for purposes of paragraph 4(b) include:

(a) joining industry organisations intended to address serious online harms; and

(b) sharing information on best practice approaches, that are relevant to the service; and

(c) working with the Commissioner to share information, intelligence, best practices and other information relevant to addressing categories of class 1A material or class 1B material that are relevant to the service; and

(d) collaborating with non-government or other organisations that facilitate the sharing of information, intelligence, best practices and other information relevant to addressing categories of class 1A or class 1B material that are relevant to the service.

25 Safety features and settings—class 1A material and class 1B material

(1) This section applies to the following services:

(a) a Tier 1 designated internet service;

(b) a Tier 2 designated internet service;

(c) an end-user managed hosting service;

(d) a high impact generative AI DIS; and

(e) a machine learning model platform service.

Minimum requirements—generally

(2) The provider of a service must:

(a) carry out an assessment of the kinds of features and settings that could be incorporated into the service to minimise the risk that class 1A material and 1B material:

(i) will be accessed by, or distributed to, end-users in Australia using the service; or

(ii) will be stored on the service; and

(b) determine, on the basis of the assessment, the most appropriate and effective features and settings for the service; and

(c) ensure that the service at all times incorporates the features and settings so determined.

Further minimum requirements for a provider of a Tier 1 designated internet service

(3) Without limiting subsection (2), the provider of a Tier 1 designated internet service must:

(a) implement measures that ensure that material can only be posted to or distributed on the service by a registered account holder; and

(b) make clear in terms and conditions, community standards, and/or acceptable use policies that an Australian child is not permitted to hold an account on the service.

(4) The provider of the service must take appropriate action to:

(a) ensure that a child in Australia who is known by the Provider to be under the age of 18 does not become an end-user of the service; and

(b) stop access to the service by a child in Australia who is known by the provider to be under the age of 18.

26 Referral of unresolved complaints to the Commissioner

(1) This section applies to the following services:
  (a) a Tier 1 designated internet service;
  (b) an end-user managed hosting service;
  (c) a high impact generative AI DIS; and
  (d) a machine learning model platform service.

(2) The provider of a service must refer complaints from individuals (including end-users) concerning the provider's non-compliance with this industry standard to the Commissioner where the complaint is not resolved within a reasonable period.

27 Responding and referring to the Commissioner

(1) This section applies to the following:
  (a) a Tier 1 designated internet service;
  (b) an end-user managed hosting service;
  (c) a high impact generative AI DIS;
  (d) a machine learning model platform service.

(2) The provider of a service must implement policies and procedures that ensure that:
  (a) it responds in a timely and appropriate manner to communications from the Commissioner about compliance with this industry standard; and
  (b) it refers unresolved complaints to the Commissioner in accordance with section 26.

28 Giving information about the Commissioner to end-users in Australia

(1) This section applies to the following:
  (a) a Tier 1 designated internet service;
  (b) a Tier 2 designated internet service;
  (c) an end-user managed hosting service;
  (d) a high impact generative AI DIS;
  (e) a machine learning model platform service.

(2) The provider of a service must ensure that information:
  (a) describing the role and functions of the Commissioner; and
  (b) describing how to make a complaint to the Commissioner about the service; and
  (c) describing the mechanisms and process required by section 29 for the service;

  is accessible to end-users of the service in Australia at all times through a dedicated location on the website for the service. The location must be "in service", that is, not on a separate webpage to the webpage for the service.

29 Mechanisms for end-users and account holders to report, and make complaints about, information on designated internet services

(1) This section applies to the following:
  (a) a Tier 1 designated internet service;
  (b) a Tier 2 designated internet service;
  (c) an end-user managed hosting service;
  (d) a high impact generative AI DIS;
  (e) a machine learning model platform service.

(2) The provider of a service must implement and make available a mechanism that:
  (a) enables end-users of the service in Australia to report, flag, or make a complaint about material accessible on the service on the basis that the material:
    (i) is class 1A material or class 1B material; and
    (ii) is in breach of an obligation under the terms of use for the service, including an obligation to comply with acceptable use policies; and
  (b) is easily accessible on or through the service, and easy to use; and
  (c) includes or is accompanied by clear instructions on how to make a report or complaint, and an overview of the reporting and complaints process.

Note: for paragraph (a) for a high impact generative AI DIS, ***material*** includes material generated (or capable of being generated) by the service.

(3) The provider of the service must ensure that the identity of a person who makes a report or a complaint using the mechanism under subsection (2) (the ***first end-user***) is not accessible, directly or indirectly, by any other end-user of the service without the express consent of the first end-user.

(4) The provider of the service must:
  (a) document its systems, processes and technologies dealing with how it responds to reports made under paragraph 29(2)(a); and
  (b) ensure that personnel responding to reports made under paragraph 29(2)(a) have appropriate training in and experience of the provider's policies and procedures for dealing with reports.

30 Action in response to end-user reports—Tier 1 designated internet services

(1) This section applies to a Tier 1 designated internet service.

(2) The provider of a service must:
  (a) take appropriate action to respond promptly to reports made by end-users under paragraph 29(2)(a); and
  (b) ensure that an end-user who makes a report concerning class 1A or class 1B materials:
    (i) is notified promptly of the outcome of the report; and
    (ii) is able to request a review by the provider of outcome under paragraph (i); and

(iii) is notified promptly of the outcome of a review under paragraph (ii).

Note 1: A report includes a request for a review of the outcome of a report under paragraph (b).

Note 2: Without limiting section 12, appropriate action may include regular reviews of the effectiveness of the measures implemented by the service provider to ensure compliance with this section.

31 Action in response to end-user reports – other designated internet service providers

(1) This section applies to the following:
- (a) a Tier 2 designated internet service;
- (b) an end-user managed hosting service;
- (c) a high impact generative AI DIS;
- (d) a machine learning model platform service.

(2) The provider of a service must take appropriate action to respond promptly to reports made by an end-user under paragraph 29(2)(a);

32 Policies and terms of use to be published

(1) This section applies to the following:
- (a) a Tier 1 designated internet service;
- (b) a Tier 2 designated internet service;
- (c) an end-user managed hosting service;
- (d) a high impact generative AI DIS.

Note: For paragraph (d), for a high impact generative AI DIS, ***material*** includes material generated (or capable of being generated) by the service.

(2) The provider of a service must publish:
- (a) the terms of use for the service, including provisions relating to acceptable use policies; and
- (b) a statement setting out the community standards applicable to the service.

(3) The publication must be accessible on the website and application (if any) for the service.

(4) The publication must:
- (a) be in plain English; and
- (b) make it clear that class 1A material is not permitted on the service and describe the broad categories of material within class 1A material; and
- (c) describe the broad categories of material within class 1B material and specify the extent to which that material is not permitted on the service, or is subject to specified restrictions.

Note: For paragraphs (b) and (c) for a high impact generative AI DIS, ***material*** includes material that is not permitted to be generated by the service.

33 Information on actions taken by the provider

(1) This section applies to a Tier 1 designated internet service.

(2) The provider of a service must publish, through a dedicated location on the website for the service, information on the:

(a) systems, processes and mechanisms implemented by the provider; and

(b) other actions taken, or to be taken, by the provider;

to reduce the risk of end-users accessing, generating or being exposed to class 1A material and class 1B material through the service.

Division 3—Reporting requirements

34 Commissioner may require risk assessments and other information

(1) This section applies to all designated internet services to which this industry standard applies.

(2) The Commissioner may, by notice to the provider of a designated internet service, require the provider to give the Commissioner any of the following documents:

(a) the most recent risk profile determination for the service;

(b) the record, as required by section 10, of the most recent risk assessment for the service;

(c) the applicable risk methodology for the most recent risk assessment for the service;

(d) the provider's development program for a specified calendar year.

Note: For development programs see section 23.

(3) The provider must give the documents to the Commissioner within the period specified in the notice.

Note: See also section 39.

35 Reports of technical feasibility of provisions of Division 2

(1) This section applies to the following:

(a) a Tier 1 designated internet service;

(b) an end-user managed hosting service;

(c) a high impact generative AI DIS.

(2) The Commissioner may, by notice to the provider of a designated internet service, require the provider to give the Commissioner a report that specifies the extent to which it is technically feasible for the provider to comply with a specified provision of Division 2.

(3) If the report discloses that it is not, or has not been, technically feasible for the provider to use a system, process or technology as required by subsection 21(2) or (4), the report must specify the alternative action required by subsection 21(6).

Note: Section 21 is about known child sexual abuse material.

(4) If the report discloses that it is not, or has not been, technically feasible for the provider to use a system, process or technology as required by subsection 22(2) or (4), the report must specify the alternative action required by subsection 22(6).

Note: Section 22 is about known pro-terror material.

(5) The report must provide justification for the conclusions in the report.

(6) The notice may require the report to be in a specified form. The provider must comply with the requirement.

(7) A report may relate to 2 or more services.

(8) The provider must give the report to the Commissioner within the period specified in the notice.

Note: See also section 39.

36 Notifying new features of designated internet services

(1) This section applies to a Tier 1 designated internet service.

(2) If the provider of a service decides to add a new feature or function to the service, the provider must notify the Commissioner of the proposed change as soon as practicable after making the decision unless the provider considers, on reasonable grounds, that the proposed change will not significantly increase the risk that the service will be used to solicit, access, generate, distribute or store class 1A material or class 1B material.

(3) If a new feature or function is added to a service, the provider of the service must notify the Commissioner of the change as soon as practicable unless the provider determines, on reasonable grounds, that the change has not significantly increased the risk that the service will be used to solicit, access, generate, distribute or store class 1A material or class 1B material.

37 Reports on outcomes of development programs

(1) This section applies to the following services:
  (a) a Tier 1 designated internet service;
  (b) an end-user managed hosting service;
  (c) a high impact generative AI DIS; and
  (d) a machine learning model platform service.

(2) The Commissioner may, by notice to the provider of a designated internet service to which section 24 applied in respect of a particular calendar year, require the provider to give the Commissioner a report that specifies:
  (a) the activities undertaken by the provider in respect of the calendar year to implement its development program; and
  (b) the outcomes of those activities in terms of enhancing online safety for end-users in Australia.

(2) The Commissioner may, by notice to the provider, require the report to be in a specified form. The provider must comply with the requirement.

(4) The provider must give the report to the Commissioner within the period specified in the notice.

Note: See also section 39.

38 Commissioner may require compliance reports

(1) This section applies to the following services:

(a) a Tier 1 designated internet service;
(b) a Tier 2 designated internet service;
(c) an enterprise DIS;
(d) a high impact generative AI DIS;
(e) a machine learning model platform service.

Minimum requirements—generally

(2) The Commissioner may, by notice, require the provider of a service to prepare and give the Commissioner a report that:
(a) specifies the steps that the provider has taken, including measures and controls the provider has implemented, to comply with applicable minimum compliance measures in this Part;
(b) includes confirmation from the provider that the steps, measures and controls are appropriate, including reasonable supporting details and evidence; and
(c) where applicable for the relevant designated internet service, such other details as specified in subsections (7) and (8).

(3) However, the Commissioner may not request a report under this section in respect of a designated internet service:
(a) at any time prior to the first anniversary of the commencement of this industry standard; and
(b) without limiting paragraph (a), more than once in any 12 month period.

(4) The notice may require the report to be in a specified form.

(5) The provider must give the report to the Commissioner within the period specified in the notice.

Note: See also section 39.

(6) A compliance report may relate to 2 or more services.

Further minimum requirements —machine learning model platform service

(7) Without limiting subsection (2), the provider of a machine learning model platform service must ensure that any report required by this section for a calendar year:
(a) specifies:
(i) the volume of child sexual exploitation material and pro-terror material identified by the provider in relation to the service in the calendar year, where it is technically feasible for the provider to identify such material; and
(ii) the number of models made available through the service during the calendar year for which it is reasonably foreseeable that the model could be used to generate CSEM or pro-terror material;
(b) specifies the way in which the details and materials under paragraph (a) (if any) were identified;

(c) includes details of the action taken by the provider in the calendar year in respect of the details and materials identified as mentioned in paragraph (a); and

(d) specifies the average number of Australian monthly active users of the service in the calendar year, and how that number was worked out.

Example: For paragraph (b): identification through reports made to the provider, hashing or through other measures and controls implemented by the provider.

Further minimum requirements—Tier 1 designated internet service, end-user managed hosting service, high impact generative AI DIS

(8) Without limiting subsection (2), the provider of a Tier 1 designated internet service, end-user managed hosting service or high impact generative AI DIS must ensure that the compliance report:

(a) specifies the volume of child sexual exploitation material and pro-terror material identified by the provider in relation to the service;

(b) specifies the manner in which the materials under paragraph (a) (if any) were identified;

(c) includes details of the action taken by the provider in respect of materials identified under paragraph (a);

(d) specifies the average number of Australian monthly active users of the service in the prior 12 month period, and how that number was worked out.

Example: For paragraph (b): identification through reports made to the provider, hashing or through other measures and controls implemented by the provider.

Further minimum requirements—end-user managed hosting service and high impact generative AI DIS

(9) Without limiting subsections (2) and (8), the provider of an end-user managed hosting service or a high impact generative AI DIS must ensure that the compliance report sets out:

(a) details of any limitations on the service or the provider to identify, assess or take action in respect of class 1A material and class 1B material; and

(b) where relevant, a description of the design and technology features of the service giving rise to the limitations under (a); and

(c) the impact of such limitations on the matters specified in paragraphs (8)(a), (b) and (c).

39 Extension of reporting periods

The Commissioner may, on application, extend the period within which a provider must give the Commissioner a report, certificate or notification under this Division, and may do so before or after the period has expired.

Part 5—Miscellaneous

40 Complaint resolution arrangements

(1) This section applies to a designated internet service if this industry standard requires the provider to make provision in respect of complaints by end-users in Australia of the service.

(2) If a complaint in relation to the service is made by an end-user, the provider must:

(a) investigate the complaint; and

(b) notify the complainant of the outcome of the investigations and the action proposed by the provider to in consequence of the investigation.

(3) Subsection (2) does not apply if:

(a) the provider believes on reasonable grounds that the complaint was frivolous or vexatious or otherwise not made in good faith; or

(b) the matter the subject of the complaint is being investigated, or has been investigated, by the Commissioner under Division 5 of Part 3 of the Act.

41 Record-keeping

(1) The section applies to all designated internet services.

(2) The provider of a service must keep records that set out the actions that the provider has taken to comply with this industry standard.

(3) The provider must keep the records for at least 2 years after the end of the calendar year during which the action was taken.