IJM Submission On:

Draft RES and DIS Industry Standards

21 December 2023

Introduction

International Justice Mission (IJM) welcomes this opportunity to provide a formal Submission on the draft Industry Standards for DIS and RES services for the Online Industry (Class 1A and Class 1B) Material under the *Online Safety Act 2021*, jointly prepared by IJM Australia and IJM's Center to End Online Sexual Exploitation of Children. We commend Australia's eSafety Commissioner for detailing measures in the Standards by which digital service providers can proactively detect and remove the most harmful online content and take greater responsibility to ensure a safer online environment.

Since 2011, IJM has worked closely with all levels of the Philippine Government, international law enforcement, community service organisations, survivor leaders, and other relevant stakeholders to combat online sexual exploitation of children (OSEC), with focus on the trafficking of children to produce first-generation child sexual exploitation material (CSEM) especially via livestreaming video. This form of child sexual abuse online, along with "self-generated" abuse in livestreams, are all live crime scenes happening on tech platforms.

To date,[1] IJM has supported 361 law enforcement operations, safeguarding 1,203 victims or at-risk individuals, leading to the arrest of 373 suspects and conviction of 216 offenders. IJM's Center to End Online Sexual Exploitation of Children protects children in the Philippines and scales the fight against this crime globally. The Center leverages and shares effective practices and models from IJM's Philippines program to enhance justice system and private sector responses to online sexual exploitation, resulting in sustainable child protection and offender accountability. IJM partners with the Philippine Internet Crimes Against Children Center (PICACC), a cooperation between Philippine and foreign law enforcement, including the Australian Federal Police.

Livestreamed child sexual abuse requires urgent attention by tech platforms because it involves repeated hands-on sexual abuse of predominantly pre-pubescent children by trusted adults in real-time as directed and paid for by foreign sex offenders. Hiding behind their screens, many Australians direct and pay for the sexual abuse of young children in livestreams on popular video chat apps.[2] One study found that 18% of online sexual exploitation cases in the Philippines were initiated by Australia-based offenders.[3] CSAM is also produced and distributed live through grooming of children directly by Australian and other offenders online. A recent study conducted by IJM in partnership with the Nottingham Rights Lab in the UK found that in 2022

[1] As of 10 December 2023.
[2] AIC (2021). For example, a study by the Australian Institute of Criminology found that 256 Australians spent more than $1.3 million over 13 years to commission and watch livestreamed sexual abuse of Filipino children. https://www.aic.gov.au/sites/default/files/2021-10/ti639_live_streaming_of_child_sexual_abuse.pdf
[3] IJM (2020) *Online Sexual Exploitation of Children in the Philippines: Analysis and Recommendations for Governments, Industry, and Civil Society.*

alone nearly half a million Filipino children were trafficked to produce new child sexual exploitation material.[4]

CSAM livestreamed in video calls allow Australian offenders to produce child sexual abuse material of children anywhere in real-time, with less digital evidence than image- or video-based CSAM distribution. Detection, reporting, and technological prevention of this type of online abuse is critical because the victims are being repeatedly abused "live." IJM's 2020 study of livestreamed child sexual abuse in the Philippines found that victims were abused on average for two years prior to intervention, in part because technology and financial sector companies failed to detect and report in real-time the crimes happening on and through their platforms.[5]

Both Industry Standards

A. Definitions of CSEM and CSAM under RES Section 6 and DIS section 6

The definitions for child sexual exploitation material (CSEM) and child sexual abuse material (CSAM) are clear and appropriate. However, live casework conducted in collaboration with the Philippine government has identified that children are forced to perform acts of bestiality in abuse material sold to offenders living in countries such as Australia. Because of this, IJM recommends that *bestiality* be included in the definition for CSEM to appropriately categorize this type of child sexual exploitation.

B. Detecting and removing known CSAM (RES section 20 and DIS section 21)

IJM is in agreement with the approach taken to make this compliance measure apply to all pre-assessed RES and Tier 1 RES providers, and not to have separate category for encrypted services.

> *Discussion Question 4: Is the technical feasibility exception in the obligation to detect and remove known child sexual abuse material and pro-terror material appropriate? How effective will this obligation be with this exception?*

IJM does not consider the technical feasibility exception appropriate with respect to the obligation to detect and remove known child sexual abuse material, as existing technological tools make detection of known CSAM feasible across all platforms, including in encrypted environments.

Professor Hany Farid who developed PhotoDNA notes "[r]ecent advances in encryption and hashing mean that technologies like PhotoDNA can operate within a service with end-to-end encryption… Another option is to implement image hashing at the point of transmission, inside the Facebook apps on users' phones—as opposed to doing it after uploading to the company's servers. This way the signature would be extracted before the image is encrypted, and then transmitted alongside the encrypted message. This would also allow a service provider like Facebook to screen for known images of abuse without fully revealing the content of the encrypted message."[6]

---

[4] IJM (2023), *Scale of Harm: Estimating the prevalence of trafficking to produce child sexual exploitation material in the Philippines* https://www.ijm.org.ph/resources

[5] IJM (2020) *Online Sexual Exploitation of Children in the Philippines: Analysis and Recommendations for Governments, Industry, and Civil Society.*

[6] https://www.wired.com/story/facebooks-encryption-makes-it-harder-to-detect-child-abuse/

Furthermore, the 2023 WeProtect Global Alliance Global Threat Assessment marks 3 ways to detect in E2EE:

- "Client-side scanning, which involves scanning messages on devices for matches or similarities to a database of illegal child sexual abuse material before the message is encrypted and sent).
- Homomorphic encryption. This is the use of a different type of encryption which allows operations to be performed without data decryption at any point).
- Intermediate secure enclaves, which decrypt the message at server level by a third party and use tools to detect child sexual abuse materials."[7]

SafeToWatch, a safety tech tool developed by UK-based SafeToNet says "files are scanned locally by WhatsApp to protect the user from harmful content being received on their device. If embedded on WhatsApp, SafeToWatch would work in exactly the same way. SafeToWatch can scan images & videos shared via the platform locally to check for the prevalence of child sexual abuse material."[8] Even the UK's Cyber Security authority has undertaken a study to demonstrate that client-side scanning within end-to-end encrypted environments is a feasible path forward to preventing child sexual abuse material distribution.[9] The UK's Internet Watch Foundation explained that "[t]hrough our collaboration with Cyacomb, we've helped to create a tool that could block images and videos of children suffering sexual abuse from being uploaded into end-to-end encrypted platforms, where it would be impossible to trace them."[10] Apple's 2021 proposal to detect CSAM on-device demonstrates that it is technologically feasible to detect and prevent CSAM in end-to-end encrypted environments at scale.[11]

A summary of available tools can be found in the accompanying document "Tech Solutions to Protect Children Online" (IJM, July 2023).

Cost considerations in assessing technical feasibility:

The technical feasibility exception allows for cost considerations to be a reason for a service provider to determine it is not technically feasible to detect and remove known CSAM; however, many of the technologies are available free of charge or for low cost.

For example, the following tools are available for no cost:

- Microsoft's PhotoDNA
- NCMEC's Hash Sharing
- Google Content Safety API and CSAI

Other tools are available for low cost. For example, the monthly cost of DragonflAI, for 500,000 active users is approximately £1200.[12] For Thorn's Safer tool, a 12-month subscription based on 1M queries per month is $30,720 USD.[13]

[7] https://www.weprotect.org/global-threat-assessment-23/
[8] https://safetonet.com/en-gb/safetowatch-end-to-end-encryption-privacy-commentary-on-wired-coverage/
[9] https://arxiv.org/abs/2207.09506
[10] https://annualreport2022.iwf.org.uk/wp-content/uploads/2023/04/IWF-Annual-Report-2022_FINAL.pdf
[11] https://www.apple.com/child-safety/pdf/CSAM_Detection_Technical_Summary.pdf
[12] Pricing | DragonflAI
[13] AWS Marketplace: Safer Essential: API-based CSAM detection built by Thorn (amazon.com)

Impact of feasibility exception on effectiveness of obligation to detect & remove

We have concerns that the technical feasibility exception would make the detection and removal obligation less effective, by providing an "out" for service providers who are not already detecting for known CSAM from adopting technological tools and fully taking responsibility to ensure that their services are not being used to access, store or distribute CSAM. Further, service providers who move to full encryption on their platforms should not be allowed to use the technical feasibility exception as a reason to not detect or remove CSAM but should be required to incorporate the capability to detect and to remove CSAM into their platform design *before* moving to full encryption. The US-based National Center for Missing and Exploited Children (NCMEC) says "[b]ased on recent disclosures from ESPs, NCMEC anticipates that widespread adoption of end-to-end encryption by reporting ESPs will begin at some point in CY 2023 and eventually may result in a loss of up to 80% of NCMEC's CyberTipline reports."[14]

Clarification on the exemption

If the exception to the obligation to detect and remove where "it is not technically feasible for the provider to do so" is maintained, there should be clear guidance that the exception applies only to the extent of the infeasibility. For example, a provider may determine that it is not technically feasible to use detection technologies on its encrypted services, but it would still be required to employ such systems, processes or technologies on parts of the service that are not encrypted. This should be clearly spelled out within the standard, not just in an explanatory document. Additionally, each provider should be required to apply for a technical feasibility exemption, with the eSafety Commissioner holding final approval.

The wording of section 7 should also ensure that the assessment of risk vs. cost should include not only the risk to the online safety of end-users in Australia, but also the risk of end-users in Australia *perpetrating* online harm. Proposed wording would be to add to the end of the section:

> **Section 7** Technical feasibility
>
> In considering whether it is or is not technically feasible for the provider of a relevant electronic 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 level of the risk to the online safety of end-users in Australia of not taking the action~~
>
> i. the level of risk to the online safety of end-users in Australia; and
> ii. the level of risk of Australian end-users perpetuating online harm
>
> of not taking action.

Discussion Question 5: Are there other examples of systems, processes and technologies that can detect, flag and/or remove known child sexual abuse material and known pro-terror material at scale, which should be highlighted in the Standards or accompanying guidance?

Below are examples of technologies that can detect, flag and remove **known CSAM** at scale, which should be highlighted in the Standards or accompany guidance:

[14] https://www.missingkids.org/content/dam/missingkids/pdfs/OJJDP-NCMEC-Transparency_2022-Calendar-Year.pdf

- Safety technology company, **SafeToNet**, has created a real-time video & image threat detection technology, SafeToWatch,[15] capable of determining whether visual data represents undesirable and illegal content such as pornography, sexually suggestive imagery, cartoon pornography, and/or CSAM. The machine-learning algorithm will hash images with harmful content and render the content harmless. SafeToNet can provide more information to the eSafety Commissioner or industry associations upon request.
- **Thorn**, a child protection technology developer, created Safer[16] to scale CSAM detection, increase content moderation efficiency, and optimise detection using advanced AI technology. It identifies known and first-generation CSAM, leveraging cryptographic, perpetual hashing and machine learning algorithms to detect CSAM at scale and disrupt its viral spread.
- Google's Content Safety API and CSAI Match[17] uses programmatic access and artificial intelligence to help platforms classify and prioritise billions of images for review. The higher the priority given by the classifier, the more likely the image contains abusive material, helping platforms prioritise human review and make their own content determinations.
- Cyacomb Safety,[18] a detection technology designed for end-to-end encryption protects personal privacy while anonymously matching and detecting known CSAM with shared user content.
- Microsoft PhotoDNA creates a unique digital signature (known as a "perceptual hash") of an image which is then compared against signatures of other photos to find copies of the same image. The database of known images is used to detect, disrupt, and report the distribution of child exploitation material. It is a privacy protective tool that cannot be used to identify a person or object in an image.
- NCMEC Hash Sharing[19] is a database held by NCMEC of confirmed CSAM hashes accessible by tech companies to detect previously categorized CSAM on their platforms.

C. Extending detection and removal requirement to include new CSAM (for both RES and DIS)

The requirement to "detect and remove" should be extended to identify and remove new and previously unrecognised material. IJM emphasizes the urgency of this expansion based on its Scale of Harm study, which revealed that nearly half a million Filipino children were trafficked to produce new child sexual exploitation material in 2022 alone.[20] The prevalence of new material underscores the real-time nature of the abuse these children endure. Focusing solely on known CSAM detection, particularly as global legislative trends move in that direction, may

[15] https://safetonet.com/safetowatch/
[16] https://safer.io/
[17] https://protectingchildren.google/tools-for-partners/
[18] https://www.cyacomb.com/company/news/2022/september/first-line-of-defence-cyacomb-launches-online-safety-software-to-combat-child-sexual-abuse-whilst-protecting-privacy/?utm_source=ActiveCampaign&utm_medium=email&utm_content=News+and+opportunities+from+across+the+Alliance&utm_campaign=September+2022+Newsletter
[19] https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata#:~:text=Hash%20Sharing&text=When%20an%20image%20or%20video,be%20identified%2C%20reported%20and%20removed
[20] https://www.ijm.org.ph/resources

inadvertently incentivise offenders to seek or produce new material to evade detection. This shift poses a significant risk, as the majority of circulating content could become new and untraceable, heightening the exploitation of vulnerable children. Therefore, a comprehensive approach that addresses both known and new CSAM is essential to effectively combat the evolving challenges posed by online child exploitation.

Technological tools currently exist to detect new CSAM

Concerns have been raised about whether there exist appropriate technological tools, systems and processes that would be able to detect first-generation CSAM, at scale and at a sufficiently high level of accuracy. Technologies and processes aimed at detecting first-generation CSAM currently exist and are being deployed on a range of services, leveraging advanced technologies like hashing, machine learning, and artificial intelligence to effectively identify and eliminate CSAM content.

Below are some examples of technological tools or actions taken by platforms in real-time to address **new and livestreamed CSAM**.

- Safety technology company, **SafeToNet**, has created a real-time video & image threat detection technology, SafeToWatch,[21] capable of determining whether visual data represents undesirable and illegal content such as pornography, sexually suggestive imagery, cartoon pornography, and/or CSAM. The machine-learning algorithm will hash images with harmful content and render the content harmless. SafeToNet can provide more information to the eSafety Commissioner or industry associations upon request.
- **Thorn**, a child protection technology developer, created Safer[22] to scale CSAM detection, increase content moderation efficiency, and optimise detection using advanced AI technology. It identifies known and first-generation CSAM, leveraging cryptographic, perpetual hashing and machine learning algorithms to detect CSAM at scale and disrupt its viral spread.
- Google's Content Safety API and CSAI Match[23] uses programmatic access and artificial intelligence to help platforms classify and prioritise billions of images for review. The higher the priority given by the classifier, the more likely the image contains abusive material, helping platforms prioritise human review and make their own content determinations.
- The social livestreaming platform, Yubo,[24] proactively screens live video to keep children safe online, implementing automated prompts to users to change behaviour and disabling violative livestreams.
- DragonflAI[25] is a prevention and disruption tool that moderates livestreams completely on-device before they are streamed to platform. It detects illegal content such as CSAM and prevents content from being uploaded.

These tools should be highlighted in guidance material accompanying the RES Standard.

[21] https://safetonet.com/safetowatch/
[22] https://safer.io/
[23] https://protectingchildren.google/tools-for-partners/
[24] https://safety.yubo.live/
[25] https://www.dragonflai.co/

Accuracy

Contrary to industry concerns, these technologies are able to detect and disrupt CSAM with high levels of accuracy. For example, a performance evaluation of the tool, SafeToWatch, found that 95.25% of all CSAM images were accurately detected by the tool, with a false positivity rate of 1.23% measured against 100,000 neutral images. Additionally, "it was observed that false positive rates in images can be suppressed as low as 0.34% while still accurately detecting 86.72% of all CSAM via confidence thresholding."[26]

A concern within industry is often the need for 99.99% accuracy to avoid there being an overload of false reports to law enforcement. However, SafeToWatch does not make reports to law enforcement and the images or video recordings are not available to SafeToWatch, as the scanning is entirely on-device. Any inaccuracies solely result in the user not being able to take the picture or video.

D. Disrupting and deterring CSAM – RES Section 22 RES and DIS Section 23(1) and (2)

IJM welcomes the obligation placed on service providers to take action to disrupt and deter end-users from using the service to solicit, create, post or disseminate CSAM and agrees with the technology-neutral approach. Prevention, disruption, and deterrence methods are crucial because they offer a proactive approach to protecting children, reducing the likelihood of them becoming victims of online sexual abuse in the first place. While detection and removal are essential to preventing further harm by stopping the distribution of illicit material, it is imperative to prioritise prevention, disruption and deterrence measures that address the root causes, discourage illicit activities and prevent the harm from taking place. Combining both detection/removal and prevention/disruption/deterrence tools, as a layered approach, can enhance the overall safety measures in place.

Under the examples given for systems, processes and technologies in RES s. 22(2)&(3) and DIS s. 23(2), the tools listed should be expanded beyond technologies aimed at content identification and removal of known CSAM. Livestreaming deterrence and disruption technologies - exemplified by tools like SafeToWatch - can play a pivotal role in preventing the live streaming of abusive content. As noted previously, technological tools currently exist that have the capability to

- detect and remove new CSAM
- detect and prevent illegal content from being created, uploaded or distributed
- deter users through automated prompts to change behaviour and disable violative livestreams

Specific examples of such tools are (see previous section for details):

- SafeToNet's SafeToWatch
- Thorn's Safer
- DragonflAI
- Cyacomb Safety

These examples can be highlighted in the guidance, as per *Questions 7 of the Discussion Paper (Are there examples of systems, processes and technologies that can disrupt and deter the use*

[26] SafeToWatch Performance Evaluation Paper - Policy Version v.1.70 Oct 2023 (1) 1.pdf

of a relevant electronic service to solicit, generate, distribute or access child sexual abuse material and pro-terror material, which should be highlighted in the guidance?).

Research also indicates the success of automated warning messages as a deterrence strategy.[27]

A multifaceted approach ensures a comprehensive strategy against online child exploitation, encompassing both proactive prevention measures and reactive content identification and removal techniques.

We recommend the addition of the following in the list of systems, processes and technologies under RES s.22(3):

(3) Without limiting subsection (2), the systems, processes and technologies may include:
(a) hashing technologies, machine learning and artificial intelligence systems that scan for known child sexual abuse material or known pro-terror material; ~~and~~
(b) systems, processes and technologies that are designed to detect key words, behavioural signals and patterns associated with child sexual abuse material; and
(c) hashing technologies, machine learning, image- and video-classifiers, and artificial intelligence systems that identifies and disrupts child sexual abuse material or pro-terror material;
(d) systems, processes and technologies that display warning messages that outline the potential risk and criminality of accessing CSAM in response to user conduct.

We also recommend that DIS s. 23(2) set out the same examples of systems, processes and technologies as in RES s.22(3), including the two additional examples noted above. The examples should be in the body of the Standard, as opposed to in a "Note".

Furthermore, we recommend including a requirement for platforms to deploy indicator detection tools that identify language or other high-risk indicators of online sexual exploitation on platforms with chat communication and/or video capabilities. From IJM's 11 years of collaborative casework experience in the Philippines, we have developed a set of both financial and tech indicators that identify potential livestreamed child sexual abuse from the Philippines to western demand-side offenders. Please see the *Tech and Financial Sector Indicators of Livestreaming OSEC* document accompanying our submission.

E. Development Programs – RES Section 23 and DIS Section 24

IJM welcomes the inclusion of an obligation on service providers to establish and implement a program of development activities and investments into systems, processes and technologies to enhance the ability of the service provider to detect and disrupt CSAM and pro-terror material on their service. The evolving nature of online threats necessitates a proactive approach and ongoing development and incorporation of advancements in technology to allow for the continuous refinement of obligations to address emerging considerations in the evolving digital landscape.

[27] Recent empirical studies on the efficacy of warning messages to deter online CSAM offending have found that warning messages dissuaded internet users from viewing 'barely legal' pornography online and sharing potentially illegal sexual images. See J. Pritchard 2022, Warning messages to prevent illegal sharing of sexual images: Results of a randomised controlled experiment (aic.gov.au)

Monthly active user threshold – RES s. 23(1) and DIS s. 24(1)

To proactively address and prevent use of platforms for accessing, distributing, creating or storing CSAM, we recommend that **all online services, regardless of size**, be required to either invest in development programs or trial and implement existing prevention technologies employed by larger platforms. This proactive stance contributes to a safer online ecosystem by mitigating risks and fostering responsible content moderation practices.

The thresholds based on monthly active user numbers, as outlined in RES s. 23(1) and DIS s. 24(1) are an appropriate basis for any differentiation in the development and investment obligation.

Enhancing ability to remove CSAM

The development program requirement includes investments and activities geared to enhancing the provider's ability to *detect and identify CSAM* and to *deter and disrupt* end-users from using the service to create, access, store or distribute CSAM [RES s. 23(3) and DIS s. 23(4)]. We recommend that the provision also include enhancing the ability of the service provider to *remove* CSAM (as well as detect it). Swift content removal is essential to prevent revictimisation of survivors and to maintain a safe online environment.

Development activities

IJM recommends adding to the list of examples of activities that could be part of a provider's development program under RES s.23(5) and DIS s. 24(7) -

> "collaborating with local and foreign law enforcement to facilitate the sharing of intelligence and other information relevant to addressing class 1A material that are relevant to the service".

IJM has noted from working on cases of livestreamed child sexual abuse with the Philippines Internet Crimes Against Children Center that service providers often become bottlenecks in the ability of law enforcement to obtain vital information to act on cases. Fostering better collaboration with foreign law enforcement agencies can expedite the process of identifying potential victims, timely intervention and the safeguarding of victims.

Investment activities

We recommend including *collaboration with survivors* under section RES s.23(6) and DIS s.24(6) as a significant investment activity. Survivor collaboration brings valuable perspectives and insights that can shape more effective strategies and interventions in combating online exploitation. Inclusion of these aspects would further strengthen the proposed development programs and investments, ensuring a more holistic and impactful approach.

F. Online safety protections should extend beyond end-users in Australia

The safety risks and impacts of internet misuse on Australian platforms are not confined to online harms to Australian end-users; many Australians are involved in exploiting and causing online harm to others outside of Australia. As referenced previously, IJM's 2020 study of livestreamed child sexual abuse in the Philippines[28] found that Australians accounted for nearly 1 in 5 offenders who engage in livestreamed sexual abuse of children in the Philippines. None of

[28] IJM (2020) *Online Sexual Exploitation of Children in the Philippines: Analysis and Recommendations for Governments, Industry, and Civil Society.*

the child victims were Australian end-users, yet online platforms available in, and used by Australians, were weaponised for that harm.

The Industry Standards should require service providers to address all aspects of online harm from the misuse of platforms based in, or accessible in, Australia. In the case of online child sexual abuse, this would include

- the direct harm to the child from the online abuse
- re-traumatisation and continuing harm to the victim from images and videos of their abuse being accessed and distributed online;
- harm to end-users from exposure to CSAM
- online harm committed by Australia-based users or through the use of Australian digital platforms, to children globally.

The definition of "online safety for Australians" in the *Online Safety Act* -

> online safety for Australians means the capacity of Australian to use social media services and electronic services in a safe manner -

focusses on the safe use of electronic services by Australians. "Capacity to use ... in a safe manner" would include not using Australian-based platforms to cause online harm (irrespective of where a potential victim resides.) We recommend that this definition of "online safety for Australians" be repeated and made explicit in the Industry Standards.

RES Section 4 further sets out the object of the Industry Standard:

> Section 4 Object of this industry standard
>
> The object of the industry is to improve the online safety for Australians in respect of class 1A material and class 1B material, including by ensuring that providers of relevant electronic 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.

Similar wording is found in section 4 of the DIS Industry Standard.

This object should be consistently referenced throughout the RES and DIS Industry Standards, highlighting that the obligations on service providers aim not only to manage and reduce risk of Australian end-users being exposed to class 1A and 1B material but also to manage the risks that Australians would solicit, generate or distribute class 1A or class 1B material through the services – *ie*. using the services to perpetrate harm. Many children are subjected to devastating abuse and harm through the use of online platforms at the hands of adult end-users, without being users of the platform themselves.

We recommend repeating and making explicit throughout both RES and DIS Industry Standards where there is reference to online safety of/for Australians, this object of preventing Australian users from soliciting, generating or distributing class 1A and class 1B material. Some examples are as follows -

Provision in draft Industry Standard	Recommended amendment
RES Standard Section 11 This Part not exhaustive	Add after "online safety for Australians" –

This Part does not prevent the provider of a relevant electronic service from taking measures, in addition to and not inconsistent with those required by this Part, to improve and promote online safety for Australians. **DIS Standard, section 11** has similar wording	and prevent Australians from using the service to perpetuate online harm.
RES Standard **Section 7** Technical feasibility In considering whether it is or is not technically feasible for the provider of a relevant electronic 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 level of the risk to the online safety of end-users in Australia of not taking the action **DIS Standard, section 7** has similar wording	Amend subsection (b) to read whether it is reasonable to expect the provider to incur that cost, having regard to ii. the level of risk to the online safety of end-users in Australia; and iii. the level of risk of Australian end-users perpetuating online harm of not taking action.
RES Standard **Section 12** What is appropriate action? (b)(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. **DIS Standard, section 12** has similar wording	Amend subsection (b)(iii) to read – 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.
Section 15(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 a law enforcement authority, or otherwise as required by law. **DIS Standard, section 15(2)** has same wording	Strike out the words ~~"in Australia~~" in subsection (b). If material affording evidence of a serious and immediate threat to life or physical safety of a person is hosted or available on a service in Australia, the service provider should be required to notify law enforcement authorities.
RES Standard **Section 17** Responding to breaches of terms of use or community standards: Class 1A material	This subsection should also reference harms inflicted by Australian users. Service providers should enforce their terms and conditions with respect to Australian users, where breach of the Terms & Conditions may

... (3) Without limiting what is appropriate action, appropriate action may include 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 provider's contractual rights under the terms of use for the service in relation to the breach. **DIS Standard, section 17(4)** has similar wording	reasonably be expected to result in online harm – regardless of whether the individual suffering harm is Australian. Amend subsection to read, after "may include exercising," – "in a way that is proportionate to the extent of online harm to Australians or caused by Australian users, that may reasonably be expected to flow from the breach …"
RES Standard **Section 27** Mechanism for end-users and account holders to report, and make complaints about, material accessible through relevant electronic services **DIS Standard, section 29** is a similar provision	The requirement for the provider of the service to provide a mechanism, tool or process that enables users to identify, flag, report or make a complaint about material accessible through the service that is in breach of the Terms of Use or community standards should be available to all users, not just Australians, where the violative material is accessible through Australian service providers.

Contact:

John Tanagho
Executive Director
IJM's Center to End Online Sexual Exploitation of Children
LinkedIn| ijm.org.ph/Center

Hiroko Sawai
Analyst, Advocacy Research
IJM Australia
[redacted] | IJM.org.au