CHILDLIGHT

Global Child Safety Institute

Authors:

Ms Zoe Lonard
Special Counsel, Norton Rose Fulbright Australia

Dr Konstantinos Kosmas Gaitis
Research Fellow (Policy & Legal Research), Childlight – Global Child Safety Institute, University of Edinburgh

Dr Mengyao Lu
Research Fellow, Childlight – Global Child Safety Institute, University of Edinburgh

Mr James Stevenson
Technology-Facilitated CSEA Data Specialist, Childlight – Global Child Safety Institute, University of Edinburgh

Professor Deborah Fry
Personal Chair of International Child Protection Research, Childlight – Global Child Safety Institute, University of Edinburgh

Legal challenges in tackling AI-generated child sexual abuse material across the 5 Eyes nations: Who is accountable according to the law?

AUSTRALIA & NEW ZEALAND

Table of Contents

Abbreviations

ACMA – Australian Communications and Media Authority

ACT – Australian Capital Territory

AI – Artificial Intelligence

CDPP – Commonwealth Director of Public Prosecutions

CPPA – Child Pornography Prevention Act

CSAM – Child Sexual Abuse Material

EEA – European Economic Area

HDCA – Harmful Digital Communications Act 2015 (NZ)

ICMEC – International Centre for Missing & Exploited Children

NSW – New South Wales

NT – Northern Territory

OECD – Organisation for Economic Co-operation and Development

OSA – Online Safety Act

QLD – Queensland

SA – South Australia

TAS – Tasmania

VIC – Victoria

WA – Western Australia

Executive Summary

This study is among the first to critically review the regulatory context of the Five Eyes nations (UK, USA, Canada, Australia and New Zealand) on the topic of accountability around child sexual abuse material (CSAM) created via generative artificial intelligence (gen-AI). This report details our findings in Australia and New Zealand. With regards to Australia, the same issues were examined on both a national as well as a state and territory level, whereas in New Zealand they were examined only on the national level, as no relevant legislation exists on a devolved basis.

It is worth considering that Australia has enacted an *Online Safety Act* (OSA) in 2021, which imposes certain duties on online service providers to protect Australians, in particular children and vulnerable adult users online. The eSafety Commissioner was established in 2015 (then the "Office of the Children's eSafety Commissioner") in Australia and serves as an independent regulator for online safety, with powers to require the removal of unlawful and seriously harmful material, implement systemic regulatory schemes and educate people around online safety risks. Under the *OSA*, there are currently enforceable codes and standards in force which apply to AI-generated CSAM with civil penalties for services that fail to comply. In particular the "Designated Internet Service Standard" applies to generative AI services, as well as model distribution services.

The Australian Government has also recently conducted consultations regarding the introduction of mandatory guardrails for AI in high-risk settings, which considers guardrails such as ensuring that generative AI training data does not contain CSAM.

No such legislation exists in New Zealand, although there are ongoing discussions and legal reform suggestions around the potential introduction of similar legislation there.

In both Australia and New Zealand, existing definitions of CSAM or similar terminology used in criminal legislation are broad enough to capture AI-generated CSAM. As a result, and despite limited case law on the matter due to the emerging character of gen-AI technologies, sentencing decisions have emerged in the Australian states of Victoria and Tasmania involving offenders who produced gen-AI CSAM. In New Zealand, no cases have yet been identified in which offenders have been sentenced for offences involving AI-generated CSAM, however, press reports suggest that offenders have been charged in relation to such material. In addition, there are reports of the New Zealand customs service seizing gen-AI CSAM, suggesting they believe they have the jurisdiction to do so. No cases have been identified in Australia or New Zealand in which AI software creators, or holders of datasets used to train AI, have been considered criminally liable in relation to the production of CSAM using their platforms or any other such charges.

In New Zealand, certain pieces of legislation (e.g. *Crimes Act 1961* and *Harmful Digital Communications Act 2015*) do not appear to apply in cases of gen-AI CSAM that portrays purely fictitious children. This is to an extent expected, as both laws require harm to be inflicted upon an identifiable natural person, and this is not the case in instances of AI-generated CSAM containing purely fictitious children.

In both Australia and New Zealand, there are no pending reforms to expand criminal accountability in relation to gen-AI CSAM to AI software creators and dataset holders. Given that the definitions of CSAM in existing criminal

legislation appear broad enough to capture AI-generated material, this is not surprising.

Recommendations

Based on these findings, the following recommendations are made for Australia and New Zealand:

Recommendation 1: Assess whether there is a need in both Australia and New Zealand to add specific offences establishing criminal accountability for AI creators, dataset holders etc.

Recommendation 2: Monitor the applicability of relevant legislation, particularly of the Australian OSA in emerging caselaw to further assess applicability

Recommendation 3: Need for policymakers and legislators in Australia and New Zealand to thoroughly assess whether civil penalty provisions should be increased in line with other jurisdictions internationally

Recommendation 4: Assess the need for introducing an Act similar to Australia's OSA in New Zealand

Introduction

Child sexual exploitation and abuse (CSEA) is considered a violation of children's rights and dignity (Ngo, 2021). A "widespread, worldwide issue of concerning magnitude" that affects both girls and boys (Simon, Luetzow & Conte, 2020: 2). CSEA may entail a series of negative effects for victims, which can impact their physical, mental or psychological health, their emotional wellbeing, social skills and interpersonal relationships, economic status, as well as vulnerability to future victimisation (Fisher et al., 2017). Within this, technology and related platforms or online environments are considered spaces which can be protective, but also pose significant risks to children's safety, increasing their vulnerability to CSEA victimisation (Simon, Luetzow & Conte, 2020). This vulnerability to victimisation is considered to be higher for children than adults (Quayle, 2016).

The rapid development of technology has led to the birth of new, immersive forms of technology, which are usually grouped under the umbrella term "eXtended Reality" (XR) (Huang, 2022). Prominent among these emerging technologies is artificial intelligence (AI), defined widely by Bahoo, Cucculelli and Qamar (2023: 1) as "the system's ability to interpret data and leverages computers and machines to enhance humans' decision-making, problem-solving capabilities, and technology-driven innovativeness". As such, and following the increasing dissemination of child sexual abuse material (CSAM) noticed across the clear and dark web, AI can prove to be a valuable tool in the efforts against CSEA by allowing the invention of detection intelligence algorithms that use deep-learning technique as methods of accurate detection of CSAM online (Lee et al., 2020; Ngo, McKeever & Thorpe, 2023). However, AI can also be misused by offenders to create CSAM with varying levels of realism that can often be hardly distinguishable from real-life material (Internet Watch Foundation, 2023). Irrespective of whether AI-created CSAM involves artificial children or children

modelled after real-life children, there is widespread concern that it can be a pathway to higher levels of CSEA offending that may include the sexual exploitation and abuse of children in real life (Internet Watch Foundation, 2023). As such, it requires a robust and clear legislative response, particularly with regards to accountability over AI-created CSAM. This call comes amidst a hotly contested debate, with some stakeholders promoting notions that CSAM created via generative AI does not hurt real children or that it may also serve to divert potential offenders from sexually exploiting and abusing real children, while others fear that generative AI-created CSAM may be the first step on a pathway towards higher offending in CSEA with real children (Internet Watch Foundation, 2023).

Based on the above, examining the existing legislative context of the Five Eyes countries, which comprise Australia, Canada, New Zealand, the United Kingdom (UK) and the United States of America (USA), becomes crucial in order to assess the readiness of their regulatory frameworks against the phenomena of AI-created CSAM and AI-facilitated child sexual exploitation and abuse. These countries have a long history of association dating back to 1956 (Weaver & Roseth, 2024). A recent study provided a review of the legal challenges that AI-generated CSAM presents in the context of Europe. Similar to the present study, this research looked at legal frameworks concerning both the creation and generation, as well as the distribution of that child sexual abuse material (Parti & Szabo, 2024). The five countries have been selected due to their democratic and open political systems, their high levels of technological advancement and literacy, as well as their progressive and advanced legislative systems, which often serve as the regulatory blueprints for other countries across the globe to model their legislation after. It also assists in forecasting the potential technological developments that have yet to be created or alternatively used in the sexual harm of children. By identifying and helping to shore up any gaps in

legislation now, it will be much easier in the future to address technology-facilitated CSEA (TF-CSEA) through the use of AI. It is especially timely as four of the five included countries are in the process of ratifying or drafting legislation to address online environment safety. The United Kingdom and Australia are in the process of implementing the respective Online Safety Acts, with Canada and the United States currently working on multiple pieces of legislation to address safety online (Ness et al., 2023). The capacity of AI-driven CSAM and child sexual abuse and exploitation will only increase with time as technology continues to develop (Parti & Szabo, 2024). It is important that legislation in countries known for combatting TF-CSEA is prepared for this. As such, it is necessary for this study to review the full breadth of legal coverage for crimes committed against children using any type of AI.

Methodology

Given that XR environments, and primarily AI, constitute a new and evolving field of technology, we anticipate gaps in legislation across the Five Eyes nations on the matter of accountability over AI-generated CSAM. To examine our research hypothesis, we conducted a legislative review of relevant laws and case law across the Five Eyes countries (USA, UK, Canada, Australia, New Zealand).

The review and analysis of the emerging pieces of legislation and caselaw was informed by the "black-letter law" approach (McConville & Chui, 2007), also known as doctrinal legal research method. Using this method, we gathered legal rules found in primary sources, such as statutes, case law, regulations, and proposed bills, and identified underlying themes or systems of application related to each source to develop a descriptive and detailed analysis of the effectiveness of existing laws, identify ambiguities and gaps, and suggest necessary legal reforms. This approach focuses on the letter of the law rather than the spirit of the law and is therefore taking a "literal approach to reading the law", as Wright (2018: 30) points out. By critically analysing primary and secondary legal sources, the aim of this approach is to restrict the number of possible outcomes, thus succinctly summarising and clarifying what the law instructs in a more systematised and narrower way than socio-legal analyses, which tend to look at the broader societal, political and policy context of legislation (Wright, 2018). The identification of themes in our legislative analysis is guided by our research hypothesis and research questions.

More specifically, we reviewed laws and case law from the Five Eyes countries on the topic of accountability with regards to generative-AI CSEA/CSAM. Legislation and case law was eligible for inclusion, if they focus on any area that intersects with accountability for CSEA/CSAM and particularly with regards to generative AI software; or if they defined concepts that are applicable and useful for

phenomena of AI-generated CSAM (e.g. case law defining the concept of obscenity). There was no defined search period, as any legislation or caselaw that can be applicable on the study topic was included. All legislative and caselaw sources were in English given that all countries studied are Anglophone nations.

To identify relevant legislations and cases across the five countries, we conducted an initial search of legal websites, such as Lexis Nexis, Practical Law, Google Scholar and Google; utilised official Government sources; and searched on local court and prosecution services' websites.

Regarding Australia and New Zealand, where cases were not identifiable via one of the above sources, we searched for and obtained copies of relevant cases via the Australasian Legal Information Institute or the New Zealand Legal Information Institute websites.

To identify potential law reforms as well as updates on the most recent cases, we conducted internet searches, consulted media sources and obtained discussion papers or reports from the relevant government agency websites. Supplementary materials, including press releases, news articles and policy reports, were identified through standard search engine queries and databases such as the Koons Family Institute/International Centre for Missing & Exploited Children (ICMEC) database and the Organisation for Economic Co-operation and Development (OECD) database.

Lastly, consulting our extensive network of experienced colleagues located in these countries crucially assisted us in locating further legislations or caselaw on the matter that we were not able to obtain via the above methods.

All identified legislations were collated and organised via Excel spreadsheets and then analysed. Traditional methods of selection process did not apply here, given that both the existence and non-existence of relevant legislative provisions

or caselaw on the studied topic have equal research value and led to important conclusions regarding the strengths and weaknesses of said legislation and regulatory frameworks of the 5 studied nations.

Data was extracted using a data extraction tool developed by the research team (https://osf.io/as83r/files/osfstorage/67851f0aaeb11fe8762f3f18). The data extracted included specific details about legislative definitions, provisions regarding accountability and other key findings relevant to the review questions. More specifically, the data extraction tool contained themes such as:

- Definitions: How reserved, devolved, federal, state, and provincial laws define terms such as “pseudo-photographs”, “indecent material”, "child pornography",[1] "obscene material," and related offenses, with a particular focus on computer-generated content.
- Accountability Provisions: Mechanisms by which individuals, platforms, and third parties are held accountable for producing, hosting, or distributing AI-generated CSAM.
- Civil Remedies: Available remedies for victims seeking compensation, particularly where AI CSAM is involved.
- Legislative Gaps: Identification of areas where legislation lacks clarity, such as the legal status of using real CSAM in AI training datasets.

This structured approach ensured a comprehensive review of current legal frameworks while highlighting areas for potential reform to meet the challenges posed by advancements in AI technology. We examined 27 pieces of legislation in Australia and in New Zealand together with 4 emerging cases.

[1] Childlight follows the Terminology Guidelines for the Protection of Children from Sexual Exploitation and Sexual Abuse. The terms ‘child abuse’, ‘child prostitution’, ‘child pornography’ and ‘rape’ are used in legal contexts.

Legislative Review: Australia & New Zealand

Introduction

With respect to Australia, we considered the research questions at the Commonwealth (national) jurisdictional level, and in each state and territory[2], namely:

(1) Australian Capital Territory (**ACT**);

(2) New South Wales (**NSW**);

(3) Northern Territory (**NT**);

(4) Queensland (**QLD**);

(5) South Australia (**SA**);

(6) Tasmania (**TAS**);

(7) Victoria (**VIC**); and

(8) Western Australia (**WA**).

With respect to New Zealand, research questions were answered at a national level as it does not have separate Commonwealth or state / regional legislation governing CSAM issues.

Existing criminal legislation and liability for AI-generated CSAM: Australia

Criminal legislation at both a Commonwealth (federal), state and territory level in Australia establish criminal liability for CSAM offending.

[2] The Commonwealth and each state/ territory of Australia has separate legislation establishing offences related to CSAM and sentencing legislation.

Commonwealth offences

The *Commonwealth Criminal Code Act 1995 (Cth) (Cth Criminal Code)* establishes criminal offences for possessing, controlling, distributing and/or obtaining CSAM outside of Australia (s273.6), and offences relating to using a postal or similar service, or carriage service for CSAM offences, and possessing, controlling, producing, supplying or obtaining CSAM for use through a postal or similar service (s471.20) or carriage service (s474.23) (among other offences).

State and territory offences

While each state and territory legislation in Australia differs in the offences created and their drafting, broadly speaking existing state and territory criminal laws establish criminal liability for:

(1) using, offering, procuring or involving a child in the production of CSAM;[3]

(2) producing CSAM;

(3) publishing, distributing, disseminating, selling or offering to sell CSAM;

(4) possessing or accessing CSAM; and/or

(5) administering a website or encouraging the use of a website to deal with CSAM.[4]

No criminal legislation we reviewed for the various Australian jurisdictions had specific offences directed at AI-generated CSAM or created specific criminal accountability for AI-generated CSAM pertaining to AI software creators, dataset holders or other such actors in the AI supply chain.

[3] ACT, NT and SA jurisdictions have a specific offence relating to this.
[4] VIC, QLD and SA jurisdictions have a specific offence relating to this.

Legislative definitions of CSAM

Commonwealth, state and territory criminal legislation in Australia each use different terminology to define CSAM. The *Criminal Code* (Cth) adopts the terminology "child abuse material," as do the jurisdictions of NSW, NT and VIC. However, the ACT, QLD, SA and WA adopt the terminology "child exploitation material". While different terminology is used, there is commonality in what types of offending are covered, insofar as it relates to CSAM material.

The legislation across the various jurisdictions in Australia generally includes "representations" or "depictions" of a person who is, or appears to be a child, engaged in an activity of a sexual nature. No criminal legislation we reviewed in Australia specifically referred to AI-generated CSAM within the definition of "child abuse material" or "child exploitation material", but some jurisdictions did expressly refer to computer generated images in their definitions (including NSW, TAS and WA). Nevertheless, our review suggests that such definitions are broad enough to cover AI-generated CSAM. Similarly, we consider that state and territory definitions of "child abuse material" or "child exploitation material" are broad and flexible enough to cover AI-generated CSAM material, adopting similar terminology of depictions, representations or descriptions. This is consistent with case law we have seen emerging in at least Tasmania and Victoria in 2024 where offenders have been charged and sentenced for AI-generated CSAM.

Existing criminal legislation and liability for AI-generated CSAM: New Zealand

In New Zealand, we considered the following legislation:

(1) *Films, Videos, and Publications Classification Act 1993 (NZ) (NZ Classification Act);*

(2) *Crimes Act 1961* (NZ Crimes Act); and

(3) *Harmful Digital Communications Act 2015* (NZ) (HDCA).

The primary legislation governing CSAM offences is the *Classification Act*, which we consider is broad enough to include offending for AI-generated CSAM. While the *NZ Crimes Act* and the *HDCA* create criminal offences pertaining to CSAM, we consider these acts are more appropriately described as being directed at material depicting or causing harm to individual children being natural, i.e. real persons. For example, our reading of the legislation is that the *HDCA* would likely apply in a scenario where CSAM involved or depicted a real child i.e. a child existing in real life, or AI was used to manipulate images of such a child to create CSAM but would not appear to apply to AI-generated CSAM depicting a non-natural, i.e. fictitious/imaginary person.

The *NZ Classification Act* contains offences (s123-124) relating to the making, copying, importing, supplying or distributing, or possessing for the purposes of supply or distribution an "objectionable publication" where the person knows or has reasonable cause to believe that the publication is objectionable. As per s131 and 131A, it is also a criminal offence to possess an objectionable publication, where a person or corporation has knowledge or reasonable cause to believe that a publication is objectionable. Punishment is imprisonment or a fine for natural persons; and for corporate bodies punishment consists of a fine. Notably, an "objectionable publication" is defined by reference to material that "describes, depicts, expresses or otherwise deals" with matters including sex, or

visual images of children who are nude or partially nude, or sexual in nature *(Classification Act, s3)*. We consider the definition of an "objectionable publication" under the *NZ Classification Act* to be broad enough to capture AI-generated CSAM, as it is sufficient that the publication "describes, depicts, expresses or otherwise deals".

Section 98AA of the *NZ Crimes Act* makes it a criminal offence punishable by imprisonment for a person to deal in people under the age of 18 for sexual exploitation, removal of body parts or engagement in forced labour. This provision of the legislation falls within a division headed "slave dealing". "Sexual exploitation" is defined to include "the taking by any means, or transmission by any means, of still or moving images of the person engaged in explicit sexual activities (whether real or simulated)" and "the taking by any means or transmission by any means, for a material benefit, of still or moving images of the person's genitalia, anus, or breasts". *(NZ Crimes Act, s98AA(3)-(6))*. While the definition of "sexual exploitation" appears intended to capture CSAM, given that the overall emphasis of s98AA of the *NZ Crimes Act* is on dealing in a *person* under the age of 18 (and acts such as selling, buying, transferring, bartering, renting, hiring, removing, receiving, transporting and importing a person) for the purposes of "sexual exploitation", we do not consider that this provision would capture offending in relation to AI-generated CSAM that does not involve a real person. Our review of the *NZ Crimes Act* suggests that there are no other provisions that would capture AI-generated CSAM.

The object of the *HDCA* is to deter, prevent, and mitigate harm caused to individuals by digital communications and provide victims of harmful digital communications with a quick and efficient means of redress. *(HDCA*, s3-4). Based on our reading of the *HDCA*, we consider that it likely has limitations in its ability to deal with AI-generated CSAM. The legislation appears to be directed at harm caused to *individuals*, being a real person, caused by a digital communication or

intimate visual recording. Therefore, if AI-generated CSAM depicted or manipulated images of an actual child, then we consider this likely to be captured by the legislation. That said, if the digital communication or visual recording was generated by AI and did not capture a real child then it is unlikely that there would be an identifiable victim or individual (being a real person) to be harmed.

Therefore, both the *NZ Crimes Act* and *HDCA* require real, identifiable children to be victimised due to the nature of offences that they regulate, and which require real people to be harmed. Thus, it is unlikely that a legislative update extending offences to fictitious children would be appropriate here.

Other legislation that provides mechanisms of accountability with regards to AI-generated CSAM: Australia

Australia's *OSA* commenced on 23 January 2022 and created a new regulatory framework to improve and promote online safety for Australians (OSA, s3). The OSA is civil legislation, with civil penalty provisions applying for certain contraventions, rather than creating criminal liability. The OSA was independently reviewed in 2024. The Review examined the operation and effectiveness and considered whether additional protections are needed to combat online harms, including those posed by emerging technologies (Minister for Communications, 2025). The Final Report of the review was tabled in Parliament in February 2025.

Table 1 below provides an overview of the key aspects of the OSA relevant to our research.

Table 1: Key aspects of the OSA relating to CSAM issues

Summary	
Basic online safety expectations	
Relevant powers	The *OSA* provides powers for the relevant Minister to determine basic online safety expectations for social media services (s13), relevant electronic services (s13A)[5] and designated internet services (s14)[6].
What is Class 1 Material?	"Class 1 Material" is defined in the OSA by reference to material that would be refused classification under the *Classification (Publications, Films and Computer Games) Act 1995* (Cth) (*OSA*, s106). On p.123, the Explanatory Memorandum to the *OSA* states that: *"…**Class 1 material would include material such as child abuse material, child sexual exploitation material… abhorrent sexual activity… the promotion of paedophile activity**, descriptions or depictions of child abuse or any other exploitative or offensive descriptions or depictions **involving a person who is, or appears to be, a child under 18 years**."*

[5] Defined in *OSA*, s13A, and includes an email service, instant messaging service, SMS service, MMS service, chat service, a service that enables end users to play online games, an electronic service specified in the legislation rules (s13A(1)). A service is exempt if none of the material on the service is accessible to, or delivered to one or more end users in Australia (s13A(2)).

[6] Defined in *OSA*, s14, and includes a service that allows end users to access material using an internet carriage service or delivers material to persons having equipment appropriate for receiving that material, where the delivery of the service is by means of an internet carriage service. A service is exempt if none of the material on the service is accessible to, or delivered to one or more end users in Australia (s14(3)).

Summary	
	The National Classification Code provides that publications (s2), films (s3), or computer games (s4) should be refused classification where they: (a) *"…describe, depict, express or otherwise deal with matters of sex … or revolting or abhorrent phenomena in such a way that they offend against the standards of morality, decency and propriety generally accepted by reasonable adults to the extent that they should not be classified; or* (b) *describe or depict in a way that is likely to cause offence to a reasonable adult, a person who is, or appears to be, a child under 18 (whether or person is engaged in sexual activity or not)."* In our view, based on the definition of RC, AI-generated CSAM is likely to be captured within the online content scheme.
Online content scheme	
Relevant powers	The Commissioner has powers to issue a: **Removal notice** (*OSA*, s109).[7] **Link deletion notice** (*OSA*, s124). **App removal notice** (*OSA*, s128).

[7] See OSA, s109 (removal notice given to the provider of a social media service, relevant electronic service or designated internet service); s110 (removal notice to a hosting service provider);

Summary	
What penalties apply?	Providers face a maximum penalty of $165,000 for failure to comply with a notice to remove harmful material. If the Commissioner is satisfied that a supplier of a social media service, relevant electronic service, a designated internet service or an internet carriage service has contravened a civil penalty provision on two or more occasions in the previous 12 months and their continued operation of that service represents a significant community safety risk, the Commissioner may make an application to the Federal Court that the supplier cease to provide the relevant service. The Federal Court may grant such an order if it is also satisfied of these criteria (ss. 156-159).
Complaints in relation to online content scheme	
Relevant powers	If a person believes that end users in Australia can access Class 1 Material provided on a particular social media service, relevant electronic service or a designated internet service, they may make a complaint to the Commissioner (s38).
Industry safety codes – social media services, app distribution services, hosting services, internet carriage services, equipment manufacturers & search engines	
Relevant powers	Part 9 of the *OSA* provides powers for industry bodies to regulate Class 1 and Class 2 harmful online material through the development of industry safety codes. There

	are currently six industry safety codes in operation[8] which apply to Class 1A and 1B material. Class 1A material is a subcategory of Class 1 material and includes child exploitation material (among other things). In terms of the code obligations across the various codes relevant parties affected as described above are obliged to (among other things): take reasonable and proactive steps to prevent access or exposure to, distribution of, and online storage of Class 1A material; take reasonable and proactive steps to limit hosting of Class 1A in Australia; consult, cooperate and collaborate with other industry participants in respect of the removal, disruption and/or restriction of Class 1A material; communicate and cooperate with eSafety in respect of matters relating to Class 1A, including complaints; provide tools and/or information to limit access and exposure to Class 1A; provide clear and effective reporting and complaints mechanisms for Class 1A; effectively respond to reports and complaints about Class 1A;

[8] https://www.esafety.gov.au/industry/codes.

Summary	
	provide clear and accessible information about Class 1A; and publish annual reports about Class 1A material and their compliance with this Code. Some of the industry codes require the parties to undertake a risk assessment.
What penalties apply?	If an industry participant fails to comply with their obligations under a code, the eSafety Commissioner may exercise enforcement powers under Part 9, Division 7 of the *OSA*. These include civil penalties (fine) and warning.
Industry standards – relevant electronic services and designated internet service	
Relevant powers	The eSafety Commissioner may by legislative instrument, determine a standard that applies to participants in a particular section of the industry, including (among other factors) if a draft code does not contain appropriate community safeguards (s145). In June 2024, the eSafety Commissioner registered industry standards that will come into effect from 22 December 2024, which apply to suppliers of relevant electronic services and designated internet services that either enable end users to communicate with one another, or allow end-users to access material using an internet carriage service, as well as gaming and dating services. The eSafety Commissioner

Summary	
	had deemed that draft codes proposed by these parties did not contain appropriate community safeguards. These standards contain obligations such as: implement appropriate systems, processes and technologies to detect and remove known CSAM; implement systems and processes to disrupt attempts by end users to use the service to solicit, generate, access, distribute or otherwise make available or store CSAM; notify as soon as practicable, if they become aware of CSAM on their service; provide mechanisms for users to make complaints; and the DIS Standard provides specific measures for high-impact generative AI, where the service has not incorporated sufficient controls to reduce the risk of generating synthetic high impact material. This may include for example, apps that "nudify" images without effective controls to prevent their application to children.
What penalties apply?	If an industry participant fails to comply with their obligations under a standard, the eSafety Commission may exercise enforcement powers under Part 9, Division 7 of the *OSA*, including warning and fine.

Other legislation that provides mechanisms of accountability with regards to AI-generated CSAM: New Zealand

New Zealand does not appear to have legislation equivalent to the *OSA*, although it has considered law reforms in this area.

Relevantly, in s4, the *HDCA* provides various remedies in relation to online content hosts:

(1) the District Court of New Zealand may, on an application, order an online content host to take down or disable public access to material that has been posted or sent (*HDCA*, s19(2));

(2) the approved agency under the *HDCA* may also lodge a notice of complaint with an online content host on behalf of a complainant (*HDCA*, s25(1));

If an online content host complies with a notice issued to it, in accordance with the requirements of the *HDCA*, no civil or criminal proceedings may be commenced in relation to it (*HDCA*, s24(1)).

Case law considering criminal liability or any other form of accountability for AI-generated CSAM offences: Australia

In Australia, there have been two recent cases in 2024 – *The King v Geale* in the Supreme Court of Tasmania and *CDPP v Smith* in the Melbourne County Court – where offenders have been convicted and sentenced for charges involving using AI to produce CSAM. The sentencing decisions show that judges were not concerned about whether existing criminal laws were equipped to deal with AI-generated CSAM. In both cases, while the relevant offending did not involve the use of real children, the Court acknowledged the harm of AI-generated CSAM in

fuelling demand for CSAM and may also lead to offenders moving toward offending involving real children.

In both cases, the Court took into account that no real children were involved in determining the sentence that should be imposed, with offences involving real children at a higher end of offending. A relevant case worth consulting is *R v Cobcroft (No 2) [2022] ACTSC 15*, which concerned accessing computer-generated images (CGI), cartoon strips, anime, hentai or comics, without any real children depicted in the material.

Case law considering criminal liability or any other form of accountability for AI-generated CSAM offences: New Zealand

In New Zealand, we are yet to identify any sentencing judgments in relation to AI generated CSAM but there have been press reports suggesting that such material is within the activities of law enforcement and the customs service.

Proposals for law reform to strengthen provisions for accountability for AI-generated CSAM: Australia

In Australia, there have been considerations at various levels of Government to strengthen regulation in relation to AI and the harms from its misuse, not necessarily limited to reforms to criminal legislation and accountability. We did not identify any specific proposals to reform criminal laws to expand accountability for AI-generated CSAM, including to other parties in the AI supply chain, such as AI software creators and dataset holders.

Table 2: Example regulatory reforms under consideration in Australia in relation to AI

Reform considered	Overview
Review of the *OSA* – November 2024[9]	In November 2024, following an independent statutory review of the *OSA*, the Federal Government announced an intention to introduce reforms to the *OSA* to introduce a Digital Duty of Care which would be imposed on technology platforms. Under this, relevant digital platforms will be required to take reasonable steps to prevent foreseeable harms on their platforms. At the time of writing this amended report, we do not believe draft legislation has been released identifying the new proposed Digital Duty of Care provision and responsibilities.
Cth – Department of Industry, Science & Resources, *Safe and Responsible AI in Australia*, September 2024 (Australian Government, 2024)	In 2024, the Australian Government opened a consultation on proposals for safe and responsible AI use in Australia. The consultation identified that Australia's current regulatory system is not fit for purpose and the distinct risks that AI poses. In September 2024, the Government introduced a proposals paper for mandatory guardrails in relation to AI, including ensuring there is adequate transparency with the public and other third parties, testing AI models to evaluate performance and compliance with guardrails, enabling human

[9] https://minister.infrastructure.gov.au/rowland/media-release/new-duty-care-obligations-platforms-will-keep-australians-safer-online

Reform considered	Overview
	oversight over AI systems, and creating corporate governance and risk management processes. The review also explored regulating "high-risk" AI settings. It found that a major issue was the expansion of hyper realistic deepfakes and the creation of synthetic CSAM. The Government has proposed that Guardrail 3 ensures that the data must also be legally obtained, ensuring that AI systems or general-purpose AI systems do not contain illegal and harmful material such as CSAM or non-consensual intimate imagery. Data sources must be disclosed. The review proposes three regulatory mechanisms which will have an impact on AI regulation of CSAM, based on either adopting existing regulatory frameworks, developing new legislative and regulatory frameworks, or developing a new cross-economy AI Act.
Cth – Department of Infrastructure, Transport, Regional Development, Communications and the Arts,	On 22 November 2023, the Minister for Communications, announced the commencement of a statutory review into the operation of the *OSA*. The focus of the current review is to assess the effectiveness of the *OSA*.

Reform considered	Overview
Statutory Review of the Online Safety Act 2021, *Issues Paper,* April 2024 (OSA Issues Paper)	Relevantly, the *OSA* Issues Paper is considering issues including: Concerns around how the tech neutrality of the *OSA* may not equip it to deal with the rapid deployment of generative AI, including the use of chatbots providing inappropriate and harmful responses to user prompts, the spread of generative AI deepfakes and the creation of synthetic CSAM (OSA Issues Paper, p.51); the review is considering whether amendments are necessary to the basic online safety expectations established by the *OSA*, to ensure that generative AI capabilities are designed and implemented with user safety in mind. whether the current civil penalty regime is sufficient. penalties under the *OSA* may fail to strike a proper balance between the various offences. For example, the maximum penalty for failing to take down illegal material such as CSAM versus failure to take

Reform considered	Overview
	down pro-terror material (OSA Issues Paper, p.33); enforceability of the civil penalty regime on individuals and platforms that are based overseas, as most online platforms Australians use are based overseas with little or no local presence (OSA Issues Paper, p.33).
NSW – Legislative Council, *Artificial Intelligence in NSW*, Report 63, July 2024 (NSW Report)	On 27 June 2023, a NSW Parliamentary Inquiry was launched into AI intelligence in NSW. A final NSW Report detailing the outcome of the inquiry was tabled in the Upper House on 25 July 2024. The NSW Report grappled with whether a specific piece of AI legislation should be introduced (similar to the approach taken in the EU and some other jurisdictions, including Canada). Ultimately, it was recommended the NSW Government work with other governments in Australia in developing an appropriate regulatory response to AI, including its safe and appropriate use. The inquiry also identified the need for a regulatory gap analysis to be completed, to identify where changes may be required to existing laws and to assist avoiding unnecessary duplication of laws. The NSW Report also observed that a separate

Reform considered	Overview
	statute governing AI "…*would be cumbersome and not keep pace with the rate and scale of change. A committee could provide continuous oversight that Parliament, its laws, and government policy respond to artificial intelligence and other emerging technologies in an iterative way.*" (NSW Report, [4.80]-[4.82], [4.84])

Proposals for law reform to strengthen provisions for accountability for AI-generated CSAM: New Zealand

We did not identify any reforms proposed in New Zealand relating to expanding criminal accountability for AI-generated CSAM, including in relation to criminal accountability for AI software creators or dataset holders. However, the New Zealand Government has considered reforms targeted at digital platforms and obligations to ensure that harmful material is not accessible on their platforms – which appear to be of a somewhat similar vein to the *OSA* introduced in Australia in 2021.

Between June 2021 and May 2024, New Zealand's Department of Internal Affairs led the Safer Online Services and Media Platforms Review. As part of the review, a discussion paper was released by the Government in June 2023. The discussion paper identified that New Zealand's main legislation dealing with such issues, including the *NZ Classification Act*, was over 30 years old and that:

(1) The current legislative regime does not have the reach and tools to deal with the online world.

(2) The system is difficult to navigate and has big gaps.

(3) Not all forms of content are covered by those bodies.

The discussion paper also set out proposals on p.22 to regulate online services and media platforms, including the options of:

(1) Developing an industry regulation model, using codes of practice for industry to achieve safety objectives.

(2) Establishing an independent regulator to approve any codes of practice, oversee their compliance and play an educative role.

(3) The new framework would continue criminal sanctions for dealing with "objectionable material".

Nonetheless, press reports from May 2024 indicate the Government ultimately abandoned the proposed reforms, citing that it was not a current priority for the New Zealand Government (The Post, 2024).

Conclusion and Recommendations

In conclusion, while different terminology is used, there is commonality in what types of offending are covered, insofar as it relates to CSAM material. No criminal legislation we reviewed in Australia specifically referred to AI-generated CSAM within the definition of "child abuse material" or "child exploitation material", but some jurisdictions did expressly refer to computer generated images in their definitions (including NSW, TAS and WA). However, our review suggests that definitions are nevertheless broad enough to cover AI-generated CSAM. Similarly, we consider that state and territory definitions of "child abuse material" or "child exploitation material" are broad and flexible enough to cover AI-generated CSAM material, adopting similar terminology of depictions, representations or descriptions. This is consistent with case law we have seen emerging in at least Tasmania and Victoria in 2024 where offenders have been charged and sentenced for AI-generated CSAM. No criminal legislation we reviewed for the various Australian jurisdictions had specific offences directed at creating criminal accountability for AI-generated CSAM pertaining to AI software creators, dataset holders or other such actors in the AI supply chain.

The primary legislation governing CSAM offences in New Zealand (*Classification Act*) is broad enough to include offending for AI-generated CSAM. On the other hand, while the *NZ Crimes Act* and the *HDCA* create criminal offences pertaining to CSAM (e.g. CSAM taking the form of still/moving images; harmful digital communications), these acts are more appropriately described as being directed at material depicting or causing harm to real-life, identifiable rather than purely artificial/synthetic children. This is to an extent expected, as both legislations require harm inflicted upon an identifiable natural person, and this is not the case in instances of AI-generated CSAM containing purely fictitious children.

The Australian *OSA* sets basic online safety expectations for social media services, relevant electronic services and designated internet services. No such legislation exists in New Zealand, although there are broader discussions regarding the introduction of a similar piece of legislation there.

In Australia, there have been limited cases to date where offenders have been sentenced for AI-generated CSAM offences. Two recent cases in 2024 – *The King v Geale* in the Supreme Court of Tasmania and *CDPP v Smith* in the Melbourne County Court – showcase that Australian legislation seems to be working adequately well against AI-generated CSAM. In New Zealand, no sentencing judgments in relation to AI-generated CSAM have been identified, but there have been press reports suggesting that such material is within the activities of law enforcement and the customs service. While it is not a criminal decision, there exists a decision of the New Zealand Film and Literature Board of Review, that determined that hentai, anime and manga could constitute an "objectionable publication" for the purposes of the NZ Classification Act.

Both countries are currently considering legal reforms to tighten regulation of digital platforms and service providers around AI-generated CSAM.

In view of the above, we recommend the following:

Recommendation 1: Assess whether there is a need in both Australia and New Zealand to add specific offences establishing criminal accountability for AI creators, dataset holders etc.

Recommendation 2: Monitor the applicability of relevant legislation, particularly of the Australian *OSA* in emerging caselaw to further assess applicability

Recommendation 3: Need for policymakers and legislators in Australia and New Zealand to thoroughly assess whether civil penalty provisions should be increased in line with other jurisdictions internationally

Recommendation 4: Assess the need for introducing an Act similar to Australia's *OSA* in New Zealand.

References

Literature Review and Methodology

Bahoo, S., Cucculelli, M., & Qamar, D. (2023). Artificial intelligence and corporate innovation: A review and research agenda. *Technological Forecasting & Social Change*, 188. https://doi.org/10.1016/j.techfore.2022.122264

Fisher, C., Goldsmith, A., Hurcombe, R., & Soares, C. (2017). *The impacts of child sexual abuse: A rapid evidence assessment*. Independent Inquiry Into Child Sexual Abuse. Retrieved from https://www.iicsa.org.uk/reports-recommendations/publications/research/impacts-csa.html

Huang, J, C. (2022). From Building Information Modeling to Extended Reality. In M. Bolpagni, R. Gavina & D. Ribeiro (Eds.), *Industry 4.0 for the Built Environment Methodologies: Technologies and Skills* (pp. 471-494). New York: Springer.

Internet Watch Foundation. (2023). *How AI is being abused to create child sexual abuse imagery*. Retrieved from https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/

Lee, H., Ermakova, T., Ververis, V., & Fabian, B. (2020). Detecting child sexual abuse material: A comprehensive survey. *Forensic Science International: Digital Investigation*, 34, 1-11. https://doi.org/10.1016/j.fsidi.2020.301022

McConville, M., & Chui, W.H. (2007). Introduction and overview. In M. McConville & W.H. Chui (Eds.), *Research Methods for Law* (pp. 1-17). Edinburgh: Edinburgh University Press.

Ness, S., Riley, C., & Bantourakis, M. (2023, September 23). Digital Governance over online safety is at risk of fragmenting. A multistakeholder approach could prevent that. World Economic Forum. Retrieved from https://www.weforum.org/stories/2023/09/its-time-for-global-alignment-on-digital-governance/

Ngo, N. (2021). Child sexual abuse violence against human dignity of children. *International Journal of Research Studies in Education*, 10(15), 97-108. https://doi.org/10.5861/ijrse.2021.a124

Ngo, N., McKeever, S., & Thorpe, C. (2023). *Determining Child Sexual Abuse Posts based on Artificial Intelligence*. Technological University Dublin. Retrieved from https://arrow.tudublin.ie/scschcomcon/392/

Parti, K., & Szabo, J. (2024). The Legal Challenges of Realistic and AI-Driven Child Sexual Abuse Material: Regulatory and Enforcement Perspectives in Europe. *Laws*, 13(6), 67. https://doi.org/10.3390/laws13060067

Quayle, E. (2016). *METHOD GUIDE 7: Researching online child sexual exploitation and abuse: Are there links between online and offline vulnerabilities?* Global Kids Online. Retrieved from http://globalkidsonline.net/wp-content/uploads/2016/05/Guide-7-Child-sexual-exploitation-and-abuse-Quayle.pdf

Simon, J., Luetzow, A., & Conte, J.R. (2020). Thirty years of the convention on the rights of the child: Developments in child sexual abuse and exploitation. *Child Abuse & Neglect*, 110, 1-8. https://doi.org/10.1016/j.chiabu.2020.104399

Weaver, J.M., & Røseth, T. (2024). The “Five Eyes” Intelligence Sharing Relationship : A Contemporary Perspective (1st ed. 2024.). London: Springer International Publishing.

Wright, L. (2018). Black-Letter Law. *LawNow Magazine*. https://www.lawnow.org/black-letter-law/

Australia & New Zealand

Attorney General Department. (2024). New criminal laws to combat sexually explicit deepfakes. Australian Ministers, Attorney General Department. https://ministers.ag.gov.au/media-centre/new-criminal-laws-combat-sexually-explicit-deepfakes-21-08-2024#:~:text=The%20Criminal%20Code%20Amendment%20(Deepfake,artificial%20intelligence%20or%20other%20technology

Australian Government, Department of Industry, Science & Resources. (2024). *Safe and Responsible AI in Australia: Proposals Paper for Introducing Mandatory Guardrails for AI in High-Risk Settings*. Retrieved from https://consult.industry.gov.au/ai-mandatory-guardrails

Australian Government, Department of Infrastructure, Transport, Regional Development, Communications and the Arts. (2024). *Statutory Review of the Online Safety Act 2021, Issues Paper (OSA Issues Paper)*. Retrieved from https://www.infrastructure.gov.au/have-your-say/statutory-review-online-safety-act-2021

CDPP v Smith [2024] VCC 1140

Classification of Computer Games and Images Act 1995 (QLD)

Classification of Films Act 1991 (QLD)

Crimes (Amount of Penalty Unit) Instrument. (2023)

Crimes (Child Exploitation Offences) Amendment Act 2023 (NZ)

Crimes (Sentencing Procedure) Act 1999 (NSW)

Crimes (Sentencing) Act 2005 (ACT)

Crimes Act 1900 (ACT)

Crimes Act 1900 (NSW)

Crimes Act 1914 (Cth)

Crimes Act 1958 (Vic)

Crimes Act 1961 (NZ)

Criminal Code Act 1889 (QLD)

Criminal Code Act 1924 (Tas)

Criminal Code Act 1983 (NT)

Criminal Code Act 1995 (Cth)

Criminal Code Act Compilation Act 1913 (WA)

Criminal Code Amendment (Deepfake Sexual Material) Bill 2024

Criminal Law Consolidation Act 1935 (SA)

Evidence Act 1929 (SA)

Films, Videos, and Publications Classification Act 1993 (NZ)

Goodenough, C. (2023, December 15). *'Get an arrest warrant': Man fined for contempt over intimate deepfake images*. Brisbane Times. https://www.brisbanetimes.com.au/national/queensland/get-an-arrest-warrant-man-fined-for-contempt-over-intimate-deepfake-images-20231215-p5ersl.html

Harmful Digital Communications Act 2015 (NZ)

McCready, S. (2024, February 18). *David Bradley Dillon Henderson: Unanderra man guilty of creating AI child abuse material*. *Illawarra Mercury*. Retrieved from https://www.illawarramercury.com.au/story/8765329/david-bradley-dillon-henderson-unanderra-man-guilty-of-creating-ai-child-abuse-material/

Minister for Communications. Report of the Online Safety Act Review released. Minister for Communications. https://minister.infrastructure.gov.au/rowland/media-release/report-online-safety-act-review-released

New Zealand Government, Department of Internal Affairs. (2023). *Safer Online Services and Media Platforms Discussion Document* (NZ Discussion Paper). Retrieved from https://www.dia.govt.nz/diawebsite.nsf/Files/online-content-regulation/$file/Safer-Online-Services-and-Media-Platforms-Discussion-Document-June-2023.pdf

New Zealand Government, Department of Internal Affairs. (2024). *Media and online content regulation*. Retrieved from https://www.dia.govt.nz/media-and-online-content-regulation#About

New Zealand Government, Department of Internal Affairs. (2024). *Safer Online Services and Media Platforms, Summary of Submissions* (NZ Summary Paper). Retrieved from https://www.dia.govt.nz/diawebsite.nsf/Files/online-content-regulation/$file/Safer-Online-Services-and-Media-Platforms-Summary-of-Submissions-V2.pdf

New Zealand Information Institute. (2021). *Video and Images (hentai style) [2021] NZFLBR 6*. Retrieved from http://www.nzlii.org/cgi-bin/sinodisp/nz/cases/NZFLBR/2021/6.html?query=title(%222021%20NZFLBR%206%22)

NSW Parliament. (2024). *Artificial Intelligence in NSW, Report 63*. Retrieved from https://www.parliament.nsw.gov.au/lcdocs/inquiries/2968/Report%20No%2063%20-%20PC%201%20-%20Artificial%20intelligence%20in%20New%20South%20Wales%20-%2025%20July%202024.pdf

Online Safety Act 2021 (Cth)

Parliament of Australia. (2021). *Explanatory Memorandum, Online Safety Bill 2021 (Cth)*. Retrieved from https://www.aph.gov.au/Parliamentary_Business/Bills_LEGislation/Bills_Search_Results/Result?bId=r6680

Parliament of Australia. (2024). *Criminal Code Amendment (Deepfake Sexual Material) Bill 2024 (Cth): Explanatory Memorandum*. Retrieved from https://www.aph.gov.au/Parliamentary_Business/Bills_LEGislation/Bills_Search_Results/Result?bId=r7205

Penalties and Sentences Act 1992 (QLD)

R v Cobcroft (No 2) [2022] ACTSC 15

R v Ramsay-Feeney [2022] ACTSC 82

Sentencing Act 1929 (SA)

Sentencing Act 1991 (Vic)

Sentencing Act 1995 (NT)

Sentencing Act 1995 (WA)

Sentencing Act 1997 (Tas)

SunLive. (2024). *AI-created child sex abuse imagery seized in NZ*. Retrieved from https://www.sunlive.co.nz/news/322779-aicreated-child-sex-abuse-imagery-seized-nz.html

The King v Kristian Troy Geale, Supreme Court of Tasmania

The Post. (2024). *Internal Affairs scraps ambitious plan to clean the internet*. Retrieved from https://www.thepost.co.nz/business/350271475/internal-affairs-scraps-ambitious-plan-clean-internet

More information

Suggested citation: Lonard, Z., Gaitis, K.K., Lu, M., Stevenson, J., & Fry, D. (2025). Legal challenges in tackling AI-generated child sexual abuse material across the 5 Eyes nations: Who is accountable according to the law? Australia and New Zealand. Edinburgh: Childlight – Global Child Safety Institute

Registered study protocol: OSF Registries | Does existing legislation on CSEA/CSAM across the Five Eyes nations and India allow for criminal liability or any other form of accountability with regards to AI-generated CSAM?

Ethics approval: University of Edinburgh, Childlight Research Ethics Sub-Committee (DELOC-KKG-0030424CL)

Advisory committee members: Professor Ben Mathews (School of Law, Queensland University of Technology), Dan Sexton (Chief Technology Officer, Internet Watch Foundation), Michael Skwarek (Manager, Codes and Standards Class 1, Industry Compliance and Enforcement, Australian eSafety Commissioner)

Funding acknowledgement: The research leading to these results has received funding from the Human Dignity Foundation under the core grant provided to Childlight – Global Child Safety Institute under the grant agreement number [INT21-01].