ROBLOX

corp.roblox.com

05 June 2026

Office of the Australian Information Commissioner (OAIC)
GPO Box 5288
Sydney NSW 2001

By email: copc@oaic.gov.au

RE: Exposure Draft of the Privacy (Children's Online Privacy) Code 2026

Roblox appreciates the opportunity to provide this submission in response to the Exposure Draft of the Privacy (Children's Online Privacy) Code 2026. As a global platform that brings millions of people together through play, the safety and wellbeing of our community – especially our younger users – is foundational to everything we build.

Crucially, protecting children on Roblox requires the responsible use of personal data. We rely on data to monitor for individuals misrepresenting their identities, detect toxic behaviour, and ensure that bad actors who breach our community guidelines are appropriately punished for their behaviour. This data feeds directly into our advanced AI algorithms, which are essential for proactive moderation and ensuring our platform remains a safe, fun place to play.

We support the OAIC's objective of uplifting privacy protections for children and young people in Australia. We also commend the OAIC's extensive consultations and the creation of age-appropriate materials to ensure children's and parents' voices directly inform this code. Through our own Teen and Parent Councils – which regularly brings together Roblox users and families to advise and seek input on safety features – Roblox shares your commitment to ensuring these important perspectives shape policy.

While we align with the overarching goals of the exposure draft, it is critical that the code does not inadvertently undermine the very safety efforts the government otherwise promotes. Restricting how platforms use data for safety, or mandating the removal of data that is essential for moderation, creates several unintended consequences.

Our primary observations and recommendations include:

- **Best interests of a child.** Comprehensive guidance is required for companies to assess the best interests of the child, as there are several competing factors in making this assessment. For example, as acknowledged in the explanatory memorandum, at times assessing the best interests of a child involves assessing the best interests of children as a community. The explanatory statement, and any subsequent guidance, should look to international

jurisdictions to provide comprehensive guidance to support companies making these assessments.

- **12-month consent termination.** Mandating that consent expires every 12 months risks creating severe consent fatigue, conditioning parents to blindly click through repetitive alerts to restore access, which may lead them to ignore critical safety updates.

- **Confirming parental responsibility.** The requirement for platforms to take reasonable steps to confirm that the person providing consent holds legal parental responsibility for the child (Section 13) may pose technical challenges. Too strictly enforcing this requirement risks creating a disproportionate privacy intrusion, as it could necessitate the collection of highly sensitive legal identity documents. We welcome an acknowledgement in the explanatory memorandum that a variety of technology can be used, as the code should allow platforms to use existing and less privacy-intrusive information and signals to reasonably confirm the relationship.

- **Right to erasure.** The proposed right to request the destruction of a child's personal information (Section 32) should explicitly exempt data retained for platform safety and moderation. If bad actors can use the provisions of the code to wipe their personal data, they can potentially evade bans and moderation to the detriment of other young Australians. Retaining this data is also strictly necessary to train our automated moderation algorithms to detect emerging harms. We thereby recommend an exemption that allows platforms to retain information for safety and moderation.

- **Publication of Privacy Impact Assessments.** While we support the principle of transparency, requiring platforms to publish a public register of complete Privacy Impact Assessments online risks exposing commercially sensitive information, and risks alerting bad actors on how to circumvent safety and moderation systems. The code should instead allow regulators to request the list of PIAs and content of the assessments privately.

We strongly recommend that the government holistically consider how to best reduce harms against children when determining how to proceed with the Children's Online Privacy Code. This evaluation must balance the safety harms that can be actively prevented through the responsible use of data against the unique security risks associated with excessive data collection.

We detail these points in the attached detailed submission.

Roblox is available to answer any questions the OAIC may have on this submission.

Sincerely,

[redacted]

Deputy General Counsel, Chief Privacy Officer

Submission:
Privacy (Children's Online Privacy) Code 2026

About Roblox

Roblox is a global platform where millions of people gather every day to imagine, create, and share experiences with each other in immersive, user-generated 3D environments. As a platform that serves a diverse and highly engaged community that includes millions of children and teens, safety is foundational to everything we build.

Recently, Roblox has implemented rigorous, age-appropriate design features. This includes the rollout of age-based account tiers—such as Roblox Kids (ages 5-8) and Roblox Select (ages 9-15)—which automatically apply high-privacy defaults, restrict specific content via maturity ratings, and limit or disable chat features based on a user's verified age. To support this, we utilize industry-leading, privacy-preserving tools like Facial Age Estimation (FAE) and ID verification to ensure our users receive the appropriate protections for their developmental stage.

Furthermore, we rely heavily on data and technology to proactively protect our community. Roblox uses robust data retention and advanced Multimodal AI algorithms, working in tandem with thousands of human moderators, to scan for and remove toxic behaviour, prevent grooming, and stop bad actors from evading platform bans. Our commitment to these data-driven safety mechanisms allows us to maintain a safe, welcoming, and fun environment for our users.

Through our dedicated Teen Council and Parent Council, we actively listen to and collaborate with the very people who use our platform, ensuring that our safety and privacy tools empower parents and protect children in practical, meaningful ways.

Detailed responses

Privacy by Default & Data Minimisation (Section 9)

Section 9 of the draft code introduces a strict high privacy by default mandate, requiring entities to automatically restrict the collection, use, or disclosure of personal information to what is "strictly necessary" to deliver the core service. While Roblox supports the philosophy of protecting younger cohorts from unnecessary collection of personal data, the current drafting of Section 9 provides barriers for our internal and legal 'safety by default' obligations.

To maintain a safe environment and actively root out bad actors, Roblox implements a sophisticated, data-driven safety framework that runs in the background of every single player session.

The "strictly necessary" standard establishes a significantly narrower and more restrictive threshold than the standard "reasonably necessary" test applied under the Australian Privacy Principles (APPs). The draft Explanatory Statement explicitly suggests that data processed for "service

improvement" or general platform safety may not meet this "strictly necessary" benchmark by default. This may create a compliance conflict with Roblox's existing legal duties under the *Online Safety Act 2021* (OSA) and the enforceable Phase 1 and Phase 2 Restructured Electronic Services (RES) Codes and Standards, plus our safety efforts that we choose to make voluntarily outside of these obligations.

The *Online Safety Act 2021* and the enforceable RES Codes legally compel platforms to achieve specific safety outcomes—such as detecting and removing pro-terror material or child exploitation—but they do not stipulate the exact technical means required to do so.

Under these regimes, Roblox is legally obligated to implement positive, proactive steps to detect, identify, and swiftly remove illegal and harmful material—including child sexual abuse material (CSAM), pro-terror content, and suicide-related material. Meeting these regulatory expectations requires the ongoing collection and processing of system-wide signals.

Roblox safety and moderation systems monitor for more than those materials that we are obligated to do under the RES codes. Restricting the definition of "strictly necessary" data to the bare minimum required to render a game instance effectively, or to that which we are legally obligated to do, penalises safety-by-design platforms that proactively monitor for systemic risks.

We recommend that the code or its explanatory materials must explicitly clarify that the collection, processing, and retention of data for the purposes of platform safety and content moderation are categorised as *strictly necessary* to provide any interactive digital service.

Best Interests of the Child (Sections 10 & 11)

Sections 10 and 11 of the draft Code require that any collection, use, or disclosure of a child's personal information must be reasonably consistent with the "best interests of the child".

There is complexity when assessing the best interests of the child. For example, while the text of the code remains focused almost exclusively on the individual child, the draft Explanatory Statement acknowledges that the "best interests" principle can apply to a group of children or children in general.

In an online environment, keeping kids safe requires platforms to make decisions based on community-wide patterns and signals. At times, obtaining and using the data of an individual child, even if not directly in that individual child's best interests, is necessary to protect the wellbeing of children as a whole.

If data collection, retention and deletion rules are applied too rigidly to an individual profile without looking at the broader ecosystem, platforms will be forced to discard the data needed to protect the wider community from emerging online harms.

Other balancing acts that companies will need to undertake are age assurance, parental controls, content restrictions and user experiences on platform.

These complexities are acknowledged in other jurisdictions, such as the United Kingdom, in providing extensive guidance and practical materials for self-assessment under their own Children's

Code. The OAIC should look to this international guidance to support Australian companies, and achieve consistency with how the best interests of a child are assessed.

The draft Australian code does not reflect these complexities, instead relying on the requirement that whatever be in “the best interests of the child”. Given the balancing assessment that companies must perform to keep platforms secure, explicit legal certainty and regulatory guidance are necessary.

The OAIC should therefore provide further guidance in the explanatory materials on how to assess the best interests of a child, and consider amending Sections 10 and 11 to ensure flexibility in assessing the "best interests of the child" (including, for example, considering the safety and digital wellbeing of the broader community of children on the service, to acknowledge that this may at times be weighted more greatly than the best interests of the individual child).

12 month consent termination (Section 15)

Section 15(4)(b) requires express consent to expire exactly 12 months after it is obtained, forcing platforms to repeatedly request permissions from parents and children on an arbitrary annual basis. This rigid, calendar-based approach does not reduce online risks, especially where privacy permissions have not changed. Instead, forcing parents into a continuous loop of re-verifying accounts for the exact same features introduces severe user friction.

Over time, this constant prompting triggers consent fatigue, causing parents to click through consent alerts to restore their child's access, which directly undermines the code's goal of encouraging meaningful engagement with privacy information. This will have flow-through effects for other updates, such as critical safety updates, that rely upon attentive engagement from users and parents.

Rather than mandating an arbitrary expiration, the OAIC should consider a transparency-led approach. We recommend amending the code to instead require annual transparency notifications that remind children and their parents of the consent previously provided, clearly explaining how their data is currently being used, and providing instructions on how to withdraw or modify that consent if they choose. This approach ensures parents remain consistently informed and in control of their child's privacy, without creating administrative friction that inevitably leads to consent fatigue.

Right to Erasure (Section 32)

Section 32 provides that entities must destroy specified personal information about a child upon request by the child or their parent/carer.

While Section 32(3) outlines exceptions to destruction (including posing a "serious threat to the life, health or safety of any individual" or for "enforcement related activities") it lacks a clear, explicit carve-out for routine platform safety and moderation.

While in some cases, the retention of such data would be allowed under the exemption under 32(3)(a), this relies on the assessment of it being a “serious threat”.

If a user engaged in toxic behaviour, this may not meet the threshold of a “serious threat to the life, health and safety of an individual” necessary to meet the exemption criteria. The child (or their parent) could use Section 32 to request the destruction of their data. If platforms are forced to delete device hashes, IP histories, or behavioral data, that bad actor can easily create a new account and return to the platform undetected. Furthermore, deleting this data degrades the historical datasets required to train AI moderation algorithms to detect emerging grooming or toxicity patterns.

We strongly recommend expanding the exceptions in Section 32(3) to explicitly include data retained for "platform safety and moderation".

Privacy Impact Assessments (Sections 38 & 39)

Section 38 requires an entity to conduct a Privacy Impact Assessment (PIA) when proposing a new service or making a significant change to data handling practices that affect children's privacy. Section 39 then requires a publicly available register of these PIAs. While transparency is a critical pillar of accountability, mandating a public-facing register of the entire library of assessments under Section 39 introduces significant unintended consequences for digital platforms.

Our safety and moderation systems rely on keeping detection criteria confidential to prevent bad actors from studying the documentation to figure out ways around platform filters. Given the use of private data in safety and moderation, there is a risk that publishing these impact assessments provide bad actors with a ‘guidebook’ on how to circumvent safety and moderation systems.

As written, the requirement to publish reviews before any new features are offered to the public also limits the ability of companies to quickly bring innovative protections to market. This would be the case even for features that were explicitly designed to enhance privacy or digital wellbeing, making responding to a newly discovered security vulnerability with an improved feature unnecessarily slow.

Further, requiring the full public exposure of a register of privacy impact assessments risks revealing highly sensitive company information, creating a distinct commercial disadvantage without providing any practical privacy benefit to families as the data collection and processing are already required to be disclosed in privacy policies.

Other jurisdictions allow regulators to privately request a comprehensive register of completed privacy impact assessments (PIAs) alongside the contents of specific assessments. Keeping this register confidential prevents malicious actors from using it to harvest insights and target unpatched vulnerabilities and supports fulsome disclosures. For example, if a hacker discovers a minor vulnerability in a widely used third-party service, they could use a public register to identify companies utilizing similar processing methods, effectively creating a roadmap to test and exploit internal systems.

OAIC should consider how to ensure that the publicly facing register neither reveals a playbook for bad actors to circumvent safety and moderation features, nor release commercially sensitive information, while still achieving the underlying goals of accountability and transparency.