Draft Children's Online Privacy Code

Integrity Institute Written Submission

About the Integrity Institute

We are a member-powered think tank and global nonprofit organization fueled by more than 600 Trust and Safety professionals across 23 countries and 80+ major platforms, all with direct experience preventing digital harms. Our credibility comes from our community's operational expertise, and we translate their insights into public-interest research, programmatic activities, policy guidance, and practical tools.

Our members have built and governed the systems that shape online experiences for billions of people. They have directly addressed critical issues such as AI governance, youth safety, targeted harassment campaigns, and hate speech by designing, deploying, and scaling solutions across nearly every major platform. This firsthand experience gives us a clear understanding of the systemic drivers of harm and the tools, data, and strategies needed to prevent or mitigate harm.

The Institute brings together one of the highest concentrations of Trust and Safety expertise, spanning industry, academia, governments, and civil society organizations. We deliver this practitioner knowledge directly to the people designing, regulating, and evaluating digital systems that shape public life. Our mission is to help build a technology ecosystem where people, societies, and democracies thrive.

Author: [redacted] Director of Research Programs at the Integrity Institute

Contributors and Experts: [redacted]

Consultation

Scope of Covered Services

The scope of "likely to be accessed by children" is potentially a form of scope creep and a potential concern, as children spend the vast majority of their time on mainstream social media, search engines, and gaming apps rather than dedicated "kids' sites." The term "primarily concerned" is also somewhat open-ended, since sites that don't necessarily target children still have a high proportion of younger users.

The OAIC should also clarify whether logged-out experiences are captured. For example, platforms like TikTok, YouTube, and many AI platforms allow users to browse, interact with, or create substantial content without an account, and it is unclear whether those platforms would be subject to the Code's obligations in that context.

The Code's scope would also benefit from potentially explicitly including smart watches, AI toys, and children's wearable devices, which are not currently mentioned but represent a significant and growing category of services likely to be accessed by children, especially in the future.

Interaction with Other Regulatory Frameworks

There is a notable inconsistency between the consent age threshold in this Code (under 15 requiring parental consent) and the social media minimum age under the Online Safety Act (under 16). The existing framework holds that under-16s are not independent enough to access social media, yet this Code treats under-15s as capable of consenting to internet-wide data collection decisions. This appears contradictory and risks creating a fragmented regulatory landscape that confuses both platforms and users.

Regarding the rollout timeline, it is important to note that many organizations struggled to determine what was "good enough" to be in compliance because the rollout guidance was vague. The OAIC should be attentive to the risk of similar outcomes here.

"Strictly Necessary" & "Best Interests of the Child"

The terms "strictly necessary" and "best interests of the child" recur throughout the Code but are not sufficiently defined, which creates significant compliance uncertainty for platforms. Entities need clear definitions of these terms to operationalize their obligations. Collecting only the minimum necessary data could also create false positives and situations where data that serves a legitimate child safety or accessibility function is inadvertently restricted.

Age Assurance: Risk-Based Approach

AI facial age estimation is gaining traction among gaming companies as a child safety and privacy tool, and is likely to become more prevalent across digital services. However, current ISO-tested facial estimation models are known to frequently miscalculate the age of transgender and nonbinary individuals or flag them as anomalies. This creates a specific and foreseeable harm where children undergoing gender-affirming hormone therapy, which can change facial bone structure, fat distribution, and skin texture, are disproportionately routed

out of low-friction verification pathways and into invasive documentation-based processes. Additionally, members of certain races, makeup, and presentation (eyeglasses) have age classifiers consistently miscalculated as well.

There is also a compounding identity document problem. Gender-diverse children often cannot yet change their name or gender on official paperwork. Requiring them to upload documentation that displays a deadname or incorrect gender marker is contrary to the "best interests of the child" as framed in the Code, and risks causing acute psychological distress, including gender dysphoria. The Code should explicitly address this.

Recommended approaches include: (a) where an AI facial estimation tool flags a user as an anomaly potentially due to medical or hormonal transition, the system must offer a non-invasive, privacy-preserving manual verification option rather than defaulting to hard identity document matching; and (b) introducing a trusted professional token mechanism — for example, a blind verification token issued by a school counsellor, therapist, or a digital mental health service — as an alternative to government identity documents for age verification purposes.

The Code should also explicitly mandate privacy-preserving age assurance mechanisms centred on strict data minimization. Specifically, the draft could require decentralized verification, such as a trusted third-party broker or device-level secure enclaves, and prohibit platforms from directly storing exact dates of birth, relying instead on age-band tokens or zero-knowledge proofs. To prevent side-channel data leaks when a user ages into a new permission tier, the Code could introduce a "cohort graduation" model: rather than changing account permissions on a user's exact birthday (which can be reverse-engineered to infer precise birth dates), users would be aged up in batches on standardised, predefined dates (bi-annually) permanently severing the link between platform permission changes and actual chronological age.

An important question is whether the Code sufficiently addresses the situation where a platform's own algorithmic systems implicitly infer that a user is a child, regardless of stated age at registration. If a platform's behavioural tracking, ad-targeting, or content recommendation engines categorize a user as a minor, the platform should be required to automatically apply child safety protections. The Code does not currently appear to address this scenario, and most of the largest social platforms use and rely on these classifiers.

It is worth noting that many of the Code's obligations rely fundamentally on age assurance functioning correctly. The OAIC should consider the market concentration risk: a small number of vendors may end up fulfilling the age verification function for a significant proportion of covered services, creating systemic dependency and single points of failure for data breaches and subsequent loss of trust. Additionally, privacy-preserving age assurance technology is still an evolving research landscape, and industry best practice is not yet established.

Universal Protections Alternative

If an entity applies the Code's protections to all users regardless of age, it avoids age assurance obligations entirely. It is worth asking what a universal online experience of this kind would look like in practice.

The Code should also be careful that the universal protections pathway cannot be used to avoid age estimation and verification altogether, in ways that prevent children from being routed into genuinely age-appropriate experiences. There is an important distinction between collecting data for safety and routing purposes versus collecting data for growth, personalisation, or algorithmic optimisation. The Code would benefit from making this distinction explicit.

"Best Interests of the Child" Standardization

The "best interests of the child" standard is used extensively throughout the Code, but its meaning is not clearly defined, and may be interpreted differently by different people and organisations.

Example: A measure such as restricting minors from making public posts represents a clear application of the best interest principle. By contrast, a content ranking system that filters age-inappropriate material for teenage users while also optimising for engagement presents a harder case: the two objectives may be in tension, and whether the overall design serves the child's best interest requires a more nuanced determination

The OAIC should provide a clearer definition, and importantly, a clearer articulation of what the standard does not mean, to ensure consistent interpretation across industry. It would be worth drawing on established human rights frameworks, including the UNCRC and UNHDR, to ground the definition.

It is also relevant to ask whether there is a national standard for age-appropriate development milestones that the Code could reference, given that what is in a child's best interests varies significantly across age groups. Developmental expectations for ages 3–5 differ substantially from those for ages 7–9 or 13–15, and a definition that does not account for this variation risks being applied inconsistently.

Finally, it may be relevant to ask from what perspective the 'best interest of the child' is taken from, as the perspective of what falls within the best interest of the child will be distinct from the parent or guardian, platform, advertiser, and the child themselves.

Parental Consent Verification

The Code's handling of parental consent verification raises a specific practical problem: to notify a child that parental consent is being sought or has been given, the platform must have a means of contacting the child directly, potentially via the child's own contact details. This creates a circularity, particularly for younger children whose accounts may be set up by parents, and the Code does not appear to address how this contact should be made. Parental control adoption at the device level is approximately 40–50%, but at the app level, it is in the single digits (2-10%). There is also a meaningful translation and education gap: it cannot be assumed that parents are sufficiently informed about data privacy to engage meaningfully with consent processes.

Consent Duration and Renewal

If consent is given but the child does not use the platform during that period, it is unclear whether consent should be considered withdrawn after a period of inactivity, or whether the 12-month clock runs regardless. It is also worth clarifying whether the annual consent renewal will form part of the initial consent notice given to users, or whether it will require a separate consent process. Will renewal require yearly notification in the same manner as the original consent?

Right to Data

The Code's requirement for 'clear and meaningful information about the logic involved' in automated decision-making risks becoming a compliance formality without a defined standard. In practice, disclosures tend toward either technical jargon that children cannot understand, or broad statements ('we use algorithms to suggest content') that satisfy the requirement in letter but not in substance. Without a floor on what 'meaningful' requires, the path of least resistance is likely to be pursued.

What would genuinely inform a child:

a. The categories of data feeding the system
b. The concrete changes the system makes to their experience
c. The actions available to them in response.

A statement such as 'if you watch a video all the way through, you'll see more like it; if you select Not Interested, you'll see fewer' conveys both a mental model and an actionable mechanism. That is the standard 'meaningful' should require. Without it, compliance will not translate into comprehension.

Right to Destruction

It is not clear from the Code whether the 30–60 day timeframe for responding to destruction requests refers to the first response or to the final resolution of the request. The consequences for non-compliance with destruction obligations are also not articulated in the draft.

In relation to data acquisition that occurs through corporate transactions, the Code does not address whether consent transfers or requires re-verification following a platform acquisition. For example, if one company acquires another, it is not clear whether existing child user consents remain valid under the new entity, or whether re-consent is required.

Parental Monitoring Notifications

The Code's notification requirements for geolocation monitoring raise a specific question about creator accounts, where parents may be managing a child's account for commercial or content creation purposes. The notification obligation in this context is unclear. At the beginning of every session? During use? Every fixed interval? A persistent banner or visual indicator may be a workable approach, but the Code should specify this rather than leaving it to platform discretion.

Privacy Impact Assessments

The Code does not articulate clearly how audits will be made publicly available. Greater transparency about the oversight and audit process — including whether audit outcomes will be made public — would strengthen accountability.

Regulatory Harmony & International Regulatory Alignment

The Codes lack provisions addressing auditability and a number of standard privacy regulatory mechanisms that practitioners familiar with GDPR would expect to see. The GDPR framework may offer useful inspiration for strengthening the Code in these areas before finalization such as areas of auditability, enforcement, and data subject rights.

General Feedback

Stakeholders have noted that several platform categories with significant proportions of child users may not be immediately visible as being within scope of the Code. By way of illustration: Chess.com has an estimated 5–15% of its 250+ million registered users under 18; Figma has an estimated 5–10%; and Canva has an estimated 20–25% of its 260+ million monthly active users. The Code would benefit from clearer guidance on how entities in these categories should assess whether they fall within scope.

It is also worth flagging the question of specialist staff welfare and oversight. The Code mandates privacy education and training for staff with regular access to children's personal information, but does not address whether additional safeguards are required for such specialists, nor whether training should incorporate law enforcement and child mental health touchpoints. The Code also does not mention what support systems should be in place for staff who are frequently exposed to high-risk content or tasked with identifying harmful patterns.