Submission on the Exposure Draft — Privacy (Children's Online Privacy) Code 2026

From mechanism-based to outcome-based regulation

[redacted]

Date: May 2026

[redacted]

TL;DR — One-Page Summary

The ask in one line. The Exposure Draft is excellent privacy drafting. With a small number of additions it could become something even more valuable: a measurable accountability regime, and a new international benchmark. The additions are additive, every duty already in the draft keeps operating exactly as written.

What the draft gets right (retain as drafted). The best-interests standard on collection and use (ss. 10–11), the strict-necessity default (s. 9), the consent and assent regime (Division 2) with its distinct 15-year threshold (s. 13), age-appropriate transparency (Division 3), mandatory Privacy Impact Assessments with a public register (ss. 38–39), child-friendly complaints (s. 36), and the right to request destruction (s. 32). These are well-drafted and unprecedented, and the recommendations below assume them as the foundation.

[redacted]

The first gap: Outcomes and evidence. The Code regulates inputs; what a platform must do, document, and disclose. It does not measure outcomes (whether children's privacy is demonstrably protected), and it does not require an independent evidence base to verify that the safeguards it mandates actually work. A platform can satisfy every section and still intrude on minors' privacy, because nothing in the Code measures the intended end state.

The second gap: Privacy as a shield. Platforms also invoke privacy in the opposite direction: refusing to build the systems that would identify minors such as age checks, age-appropriate defaults, etc., on the grounds that doing so would mean collecting data about them. The fix is to confirm that data-minimisation does not extend to processing used solely to protect children and firewalling from commercial use.

The five recommendations:

1. **Define key terms:** give the Code's existing duties determinate, enforceable meaning. Closes loopholes that platforms often use to evade the spirit of regulations (Code amendment; commences with the Code.)
2. **Constructive knowledge:** a platform is taken to know a user is a child when its own commercial systems already treat the user as one. (Code amendment; commences with the Code.)
3. **Outcome-based standards:** measure what results on the child's side of the interface, not merely what the platform documents. (Code amendment; phased over 24 months.)
4. **Independent evidence base:** secure-compute ("clean room") data access for accredited researchers, pre-registration, replication, and a public registry. (Code amendment + minor companion; phased over 36 months.)
5. **Audit architecture:** a researcher safe harbour, anti-evasion rules, and standardised evaluation methodologies that make independent measurement possible. (Online Safety Act 2021 amendment; phased over 36 months.)

If the drafters adopt only one. Recommendation 2 (constructive knowledge). It closes the structural loophole that lets platforms disclaim awareness of the minors their own systems already target, and it commences with the Code.

A note on scope. Recommendations 1 and 2 sit within the existing Privacy Act authority and apply universally. Recommendation 3 applies to services above a scale threshold. Recommendations 4 and 5 reach into territory shared with the Online Safety Act 2021 and the eSafety Commissioner, and are designed to be delivered partly through companion provisions — a boundary the submission is explicit about throughout.

Executive Summary

The Exposure Draft of the Privacy (Children's Online Privacy) Code 2026 is substantial and careful work. Its core architecture, the best-interests standard (ss. 10–11), strict-necessity defaults (s. 9), the consent and assent regime (Division 2), age-appropriate transparency (Division 3), Privacy Impact Assessments with a public register (ss. 38–39), child-friendly

complaints (s. 36), and the right to request destruction (s. 32), is a meaningful and unprecedented advance on the Australian Privacy Principles and on comparable international instruments. These provisions should be retained substantially as drafted; the recommendations below assume them as their foundation.

A few suggestions could make the Code stronger. Two gaps run through the draft, one in what it measures, another one in how it can be evaded, and both can be closed additively, without disturbing any duty already written.

The first gap: process without outcomes or evidence. The Code is largely a set of procedural duties such as notices, consents, registers, training, self-assessments. It regulates the inputs to platform conduct but does not demand a measurable child-safety result (the OUTCOMES), nor an independent evidence base to verify that the safeguards it mandates actually protect children (the EVIDENCE BASE). A platform can satisfy every section and still intrude on minors' privacy, because nothing in the Code measures the intended end state: a demonstrable limitation for collection, storage, and usage of private data of Australian children and teens.

The second gap: privacy invoked as a shield. Compliance regimes are gamed in two opposite directions. In the first, an entity processes children's data for commercial ends and calls it "necessary to provide the service." In the second case, less obvious, equally corrosive, an entity declines the processing that would protect children and invokes data-minimisation as the excuse: because it is restricted from holding data about minors, it argues, it cannot build the systems that identify them, check ages, remove underage users, or place a child into an age-appropriate experience. The argument is circular by design: every protection in the Code attaches only once a user is known to be a child, so a platform that claims it is forbidden from determining who is a child insulates itself from the entire regime. The fix is to state what data-minimisation always meant: restrictions do not apply to processing undertaken solely for a safety, security, integrity or compliance purpose, whether mandated or voluntary, provided that data is firewalled from commercial use. The "strictly necessary" definition (R1) already does this through limbs (b), (c) and (j), with the same firewall at s. 8A(4). The Code should be written holistically to close both routes, because shutting one while leaving the other open merely redirects the evasion.

Both gaps are addressed additively. The process duties already drafted keep operating; the outcome standards sit alongside them as the empirical accountability mechanism the input duties lack. Together they form a complete loop: platforms cannot feign ignorance of minors' presence, the standards are measurable, an independent evidence base verifies whether they are met, and the audit architecture makes that verification operationally possible.

The five recommendations

1. Defining key terms, including strictly necessary, best interests of the child, and likely to be accessed by children, so that the Code's existing duties carry determinate, enforceable meaning rather than meanings platforms can argue away. This could also include defining terms in ways that separate requirements depending on use cases for the data, such that nonprofits/services for public goods like government

services or educational ones, can be protected alongside safety-related data collection and processing of commercial apps.

2. A constructive-knowledge rule (a new s. 8A) providing that a platform is taken to know a user is a child when its own commercial systems already treat the user as one, closing the wilful-ignorance loophole;
3. Outcome-based standards layered onto the existing process duties, measuring what actually results on the child's side of the interface (for example, caps on the collection of personal data from children for algorithmic optimization), not merely what the platform documents;
4. An independent evidence-base mandate requiring safe "compute enclave" data access for accredited researchers, pre-registration of safeguard claims, replication, and a public registry; and
5. An audit architecture (researcher safe harbour, standardised evaluation methodologies, anti-evasion rules) that makes independent measurement possible.

The most impactful changes

Among the recommendations that follow, these would most change platform conduct and are the submission's highest-leverage proposals:

1. **Constructive knowledge (R2):** a platform is taken to know a user is a child whenever its own commercial systems already target or treat the user as one, ending the practice of disclaiming knowledge a platform plainly holds.
2. **Ad-funding is not a service function (R1, "strictly necessary"):** because the revenue model is excluded from what is necessary to provide the service, profiling-based recommendation for children cannot be a default and must instead pass through consent and a best-interests assessment. Similarly personalization algorithms should not count as "strictly necessary". Apps have built logged out versions (e.g., YouTube, TikTok have browser versions of their apps that are not requiring log in).
3. **The burden of proof reversed (R1, "best interests"):** the platform must demonstrate that its handling of a child's data is consistent with the child's best interests, rather than the regulator having to prove it is not.
4. **A ceiling on undetected minors (R3):** the proportion of a service's active users who are minors the entity has failed to identify becomes a measurable, enforceable standard.
5. **An independent evidence base (R4/R5):** secure-compute access for accredited researchers, plus a safe harbour, so that safeguard claims can be tested rather than self-certified.
6. **Safety processing is never barred by minimisation (R1, "strictly necessary"):** a platform cannot invoke data-minimisation to avoid building age-assurance or

protective systems — such processing is expressly preserved by limbs (b)/(c) and firewalled from commercial reuse by (j).

Each recommendation includes discussion of the principal objections it will face.

Scope and allocation. The submission is careful about which recommendations fit within the Code's authority under s. 26GC of the Privacy Act 1988 and which sit more naturally in companion legislation. Recommendations 1, 2, and 3 are proposed as Code amendments and operate within the existing Privacy Act authority on data-processing duties for the protection of children. Recommendation 4 requires minor companion provisions, most cleanly in the Online Safety Act 2021, to deliver the researcher safe harbour, but the institutional architecture it establishes can be set up under the Code's authority. Recommendation 5 (the audit architecture) likewise sits in companion legislation: the researcher safe harbour from civil and criminal liability requires a statutory basis the Privacy Act's rule-making power cannot supply, and is most cleanly located in the Online Safety Act 2021, while the Commissioner's power to publish standardised evaluation methodologies can be exercised under existing guidance powers. The further provisions in Part 6 sit in the broader regulatory architecture and are flagged for the drafters' consideration without being proposed as Code amendments. A scope table at the end of this Executive Summary sorts each recommendation by instrument.

Minimum viable adoption. The five recommendations are interlocking but not all-or-nothing. If the drafters adopt only one, the highest-leverage choice is Recommendation 2 (constructive knowledge), which closes the structural loophole that currently allows platforms to disclaim awareness of minor users on their services. If the drafters adopt two, Recommendation 3 (outcome standards) should follow, providing the measurable accountability standard that the existing process duties lack. Recommendations 4 and 5 can follow in subsequent iterations of the Code or in companion legislation.

Phasing. Recommendation 1 (defining key terms) commences with the Code: because the definitions clarify concepts already in the Code rather than introducing new obligations, they take effect immediately on commencement and require no phase-in. Recommendation 2 likewise commences with the Code. Recommendation 3 should phase in over 24 months from commencement, with the first 12 months used to establish baseline amplification measurements and to set initial thresholds in consultation with accredited researchers, and the second 12 months used for compliance ramp-up. Recommendations 4 and 5 should stand up over 36 months: accreditation framework and methodology standards in months 1–12, secure-compute data access protocols and safe-harbour provisions in months 12–24, full operational evidence-base and audit regime in months 24–36.

Proportionality. The recommendations are calibrated to platforms operating at the scale of major social media services, conversational AI services, and recommendation-driven content platforms. Full application to small Australian-built services with limited compliance capacity would be disproportionate and is not intended. The submission proposes that outcome standards under Recommendation 3 apply to services with minor end-user populations above a defined threshold, with simpler obligations for smaller

services. Recommendations 1 (defining key terms) and 2 (constructive knowledge) should apply universally. Neither imposes a new compliance burden that proportionality concerns are designed to filter: Recommendation 1 only fixes the meaning of duties the Code already contains, and Recommendation 2 only prevents a platform disclaiming knowledge it already holds. A small service is no more entitled to an indeterminate definition, or to wilful blindness, than a large one.

Scope Table

The following table summarises the instrument allocation, applicability, and proposed commencement for each recommendation.

Recommendation	Instrument	Applicability	Commencement
R1. Defining key terms	Code amendment (s. 4 & operatives)	All entities bound by the Code	With Code commencement
R2. Constructive knowledge	Code amendment (s. 8A)	All entities bound by the Code	With Code commencement
R3. Outcome standards	Code amendment (new Division)	Services above scale threshold	Phased over 24 months
R4. Evidence-base mandate	Code amendment (expanding s. 38) + minor companion	Services subject to R2	Phased over 36 months
R5. Audit architecture	Online Safety Act 2021 amendment	Services subject to R2	Phased over 36 months
Part 6 provisions	Broader regulatory architecture	Per instrument	Flagged, not proposed

Alignment with the Privacy Act Review. Several of the recommendations in this submission align with proposals advanced in the Attorney-General's Privacy Act Review Report (2023), including the recommendations on children's privacy, the proposed strengthening of accountability mechanisms, and the proposed introduction of a fair-and-reasonable test for personal-information handling. Where the recommendations in this submission go further than the Privacy Act Review, they do so within the children's-specific context for which the Code is the appropriate vehicle.

Recommendation 1: Define key terms clearly to prevent evasion

Much of the Code's force depends on terms that are presently left undefined. The draft defines almost nothing of its own — it supplies "age appropriate" and "person with parental responsibility" and borrows the rest from s. 6 of the Privacy Act and from the Online Safety Act 2021. The result is that load-bearing concepts — "strictly necessary," "best interests of the child," "likely to be accessed by children," "destroy" — carry whatever meaning a regulated entity can argue for. Each definition proposed here is built to survive a specific, foreseeable evasion. The full proposed drafting for all ten terms, with a critical note

on the evasion each clause is designed to defeat, is set out at Appendix A; the table below summarises where each provision belongs and the purpose it serves.

Placement and purpose summary

Term	Instrument	Purpose
strictly necessary	s. 4 definition	Sets the minimisation gate for the whole Code: by fixing what "necessary" means, it stops platforms treating ad-funding or engagement as a service function, and decides what processing survives the s. 9 default. It also answers the second gap (privacy as a shield): limbs (b)/(c) confirm that safety, security, integrity and compliance processing survives the minimisation default, while (j) firewalls it from commercial reuse.
best interests of the child	s. 4 signpost + new s. 10A	Converts an abstract standard into an auditable test via mandatory factors, a commercial-conflict tiebreaker, and a reversed burden, so the platform must demonstrate consistency rather than assert it.
likely to be accessed by children	s. 4 signpost + new s. 7A	Operates the Code's on/off switch through a rebuttable presumption of coverage: a service is in scope unless the entity proves otherwise on evidence.
service / the entity's service	s. 4 definition	Anchors "strictly necessary to provide the service" by excluding the business model, defeating the claim that ad-targeting data is necessary because the service is ad-funded.

Term	Instrument	Purpose
personal information / collects	s. 4 / new s. 4A clarification	Confirms, within existing Act authority, that inferred and derived data about a child is regulated and that generating it is "collection" — the precondition for the constructive-knowledge and amplification standards.
destroy	s. 4 definition (+ subsections)	Gives the s. 32 destruction right real reach by extending it to backups, derived profiles, and model-incorporated data, closing the "deleted the row, kept the profile" gap.
reasonable steps	new s. 8(2A)–(2C)	Deliberately left open rather than fixed, with a risk-proportionate floor and an anti-evasion limb, so verification scales with harm without burdening low-risk services.
high-risk content category	s. 4 definition + delegated instrument	Lists the harm categories the amplification limits attach to, with a delegated-instrument mechanism so the list updates without re-opening the Code. (See the authority note in Appendix A: this concept has no anchor in the privacy code and is best delivered via the Online Safety Act.)
safeguard (safety feature)	s. 4 definition	Captures safety features both by function and by the platform's own claims, so a platform cannot escape efficacy testing simply by declining to call a feature a "safeguard."
automated system / commercial purpose	s. 4 definitions	Supply the load-bearing terms for the

Term	Instrument	Purpose
		constructive-knowledge rule, defining when an inference is commercial (and so triggers knowledge) while sheltering safety and integrity systems.

Note. The full proposed drafting for Definitions 1–10, with the critical notes and residual soft-spots for each, is set out at Appendix A. It is placed there so that the Executive Summary and the substantive recommendations read as a continuous argument; the definitional machinery, while essential, is reference material the drafters will return to clause-by-clause.

Best interest vs. parental control: when the parent is the privacy risk

The Code's consent and assent architecture (Division 2) assumes that a person with parental responsibility acts as the child's privacy guardian. A class of cases inverts that assumption, so that the consenting adult is the source of the exposure. Parent-managed "kidfluencer" accounts are the clearest example: because platforms bar under-13s from holding their own accounts, a parent opens and runs an account featuring a young child, monetises it through subscriptions and brand deals, and in doing so publishes identifiable images of the child to a predominantly adult, predominantly male audience. A year-long New York Times investigation analysing more than 5,000 such mother-managed accounts found that "suggestive" posts drew the most engagement from male followers and that the audiences included many men who openly admitted a sexual attraction to children; a mother in Melbourne described wanting to stop but feeling unable to because her child's following had grown too large to walk away from.[2] The injury here is a privacy injury at its core — the collection, publication and commercial exploitation of a child's personal information and image — and it is not cured by parental consent, because in these cases parental consent is the very mechanism of exposure.

The fix requires no new doctrine, only that the Code make explicit what its own structure already implies: best interests is a duty owed by the entity, so parental consent is a gate, not a shield. Section 10A should provide that the entity's best-interests obligation applies regardless of parental consent, and that where a parent's decision conflicts with the child's best interests the latter prevails — a parent-conflict tiebreaker mirroring the commercial-conflict tiebreaker in s. 10A(2); the assessment should expressly weigh the reasonably foreseeable risk that publishing a child's images to a general or adult audience exposes the child to harm, whoever manages the account. The s. 32 destruction right should likewise survive parental consent, so that a child — in particular an older teen reclaiming control under a graduated framework — can erase material a parent published years

[2] Jennifer Valentino-DeVries and Michael H. Keller, "A Marketplace of Girl Influencers Managed by Moms and Stalked by Men," *The New York Times*, 22 February 2024, https://www.nytimes.com/2024/02/22/us/instagram-child-influencers.html.

earlier. These protections should apply universally, alongside Recommendations 1 and 2 and outside the Recommendation 3 scale threshold: they fix the application of a duty the Code already imposes rather than adding a new measurement burden, and the harm in fact concentrates on the smaller, less-scrutinised child-creator, pageant and dance applications where a parent is the consenting adult. The monetisation dimension — disabling direct monetisation of content heavily featuring a minor, and mandatory financial-trust arrangements — remains a matter for the Online Safety Act and is addressed in Part 6. The proposed drafting for s. 10A(5)–(7) and the companion s. 32 clarifier is set out at Appendix A (Definitions 2 and 6).

Recommendation 2: Constructive Knowledge

The problem

Section 8 requires entities to take "reasonable steps" to ascertain age. It does not address when an entity's own automated systems (recommendation engines, ad-targeting systems) have already inferred that a user is a child, while formal compliance systems claim no "actual knowledge." This is the central evasion strategy of the contemporary platform economy.

The recommendation

Insert a new provision establishing that an entity is deemed to know an end-user is a child if its automated systems classify, target, or treat the user (either explicitly or implicitly) as a minor for commercial purposes.

Proposed drafting

Section 8A Constructive knowledge of minor status

(1) For the purposes of this Code, an entity is taken to know that an end-user of the entity's service is a child if any automated system operated by or on behalf of the entity, and used for a commercial purpose (including advertising targeting, content recommendation, commercial user segmentation, growth optimisation, or engagement optimisation), classifies, targets, or treats the end-user as a minor, or as falling within a cohort defined wholly or partly by minor status.

(2) Subsection (1) applies regardless of whether the entity has independently verified the end-user's age through any other means.

(3) An entity may not rely on the absence of formal age-verification to defeat the application of this Code in respect of an end-user to whom subsection (1) applies.

(4) Subsection (1) does not apply where the automated system that has classified the end-user as a minor is a system whose primary purpose is to enable the entity to comply with this Code, the Online Safety Act 2021, or to protect the security, safety, and/or integrity of the service. The classifications and outputs of such a safety or integrity system must be firewalled and strictly prohibited from use for any of the commercial purposes listed in subsection (1).

Addressing the proxy behaviour and perverse-incentive objections

A predictable objection is that subsection (1) would punish platforms for building robust age-inference systems specifically to protect children. This is the function of subsection (4): classifications produced by a safety or integrity-purposed system that is firewalled from commercial use do not trigger constructive knowledge. The penalty attaches to the commercial exploitation of the inference, not the creation of the classifier itself.

Furthermore, establishing what constitutes a "minor cohort" based on proxy behaviours (e.g., listening to an artist with a predominantly pre-teen audience vs. listening to a mainstream pop star) will require regulatory definition. The Commissioner should be empowered to issue guidance on the behavioural clustering thresholds that trigger constructive knowledge.

Recommendation 3: Outcome-Based Standards

The problem

The Code's core mechanisms — strict-necessity defaults (s. 9), best-interests consistency for collection and use (ss. 10–11), informed and specific consent/assent (Division 2), age-appropriate transparency (Division 3), data destruction (s. 32), and Privacy Impact Assessments (ss. 38–39) — are all input duties. They specify what a platform must do, document, or disclose. They do not specify what must result on the child's side of the interface (e.g., no child data retained or used in model training, no long-term monitoring of child-related behaviours).

This creates predictable failure modes:

- **Compliance theatre:** platforms staff up privacy teams and meet every documentary obligation. The empirical experience of children on the service is unchanged, because none of those outcomes is regulated.
- **Self-certification bias:** PIAs under s. 38 predict harm; they do not measure it.
- **Unverifiable rights (assent and destruction):** under s. 20, a child under 15 must "assent" to features. Without anti-evasion rules, platforms will use frictionless UI design (dark patterns) to secure this assent. Parent-child linking is regulated, but the law does not mandate testing of whether that mechanism actually works. Similarly, s. 32 establishes a vital right to data destruction. Without researcher or government access to platform data structures (via clean rooms or secure enclaves), it is impossible to verify whether platforms are destroying data across all backend systems, or merely hiding it from the UI.
- **Wilful ignorance:** s. 8(1) requires only "reasonable steps" to ascertain age. This replicates the structural weakness of US COPPA. Platforms deliberately avoid acting on the behavioural and inferential data they already possess about user age, because acquiring formal knowledge triggers obligations.

Mechanism-based regulation dates badly. A rule written for the recommender algorithms of 2026 is partially obsolete by 2028. It is worth being precise about the starting point: the draft Code is presently silent on algorithmic delivery altogether — no provision addresses

recommendation, ranking, amplification, or feed design, and the only place automated systems appear is a transparency right (s. 28(2)(g)) to be told that automated decision-making or profiling is occurring. The recommendation below therefore introduces the concept rather than refining an existing one. An outcome rule — e.g., the algorithmic lift of defined harmful content to minor accounts shall not exceed X% over an organic baseline — remains operative regardless of technological evolution.

The principle

The Code should regulate the end state on the child's side of the interface. Outcome standards are routine in workplace safety and financial conduct; their absence here is an anomaly inherited from legacy privacy frameworks.

Three concrete outcome standards

1. Safeguard-efficacy thresholds. Where a platform deploys a safeguard, the feature must achieve defined efficacy on independent evaluation. The safeguards that map directly to the Code's own provisions — and therefore make the most defensible efficacy targets — are age assurance (s. 8); the parental control and monitoring mechanisms recognised in s. 33; and the consent and withdrawal mechanisms the Code already requires to be "clear, simple and easily accessible" and free of manipulation (ss. 14, 17, 19, 21). Efficacy testing of the last category means testing whether consent and withdrawal flows are genuinely usable or are covertly dark-patterned. Content-safety, grooming-detection and child-sexual-exploitation-detection systems are also safeguards, but because they belong to the Online Safety Act regime they are best cross-referenced rather than set as the Code's headline efficacy targets. A safeguard that achieves a measured efficacy below a defined floor is not a compliant safeguard.

2. Proportional undetected-minor ceilings. Undetected-minor ceiling (flat cap). Section 8 should be supplemented by a measurable outcome standard: a single, flat ceiling on the proportion of a service's active users who are minors the entity has failed to identify as such. Rather than mandating a recall-accuracy floor for age-inference systems which would push platforms toward affirmatively verifying all users, including adults, the Code caps the undetected-minor rate at a fixed percentage that applies uniformly to every service within scope (the figure to be specified in the Code and set, during the phase-in, in consultation with accredited researchers). A minor whom the entity's own systems already treat as a minor (see s. 8A) is not, for this purpose, undetected. The true minor population against which the rate is measured is estimated through the independent measurement architecture in Recommendation 4.

Why flat rather than proportional. A flat cap is deliberately chosen over a ceiling that scales with each service's estimated minor-attractiveness. A proportional ceiling would reintroduce exactly the kind of contestable, platform-argued variable the rest of this submission works to remove: the entity would litigate which tier it belongs in, and the regulator would carry the burden of proving a service is "attractive to minors" before any ceiling bit. A single figure is simpler to administer, gives every service the same clear target, and cannot be gamed by disputing one's tier. Risk-calibration is not lost — it moves to the

scope gate rather than the standard: Recommendation 3 applies only to services above the minor-population scale threshold, so the services held to the cap already have a meaningful minor presence. And because the numerator is undetected minors as a share of *all* active users, the flat cap naturally bites hardest on services that combine a large minor presence with weak detection — the intended target — while being near-trivially satisfied by genuinely adult-oriented or B2B services. The proportional tiering was therefore doing less work than it appeared.

This remains a genuine outcome standard: it measures the result — minors slipping through undetected — rather than a model-internal input; it does not force verification of adults; and it is agnostic to the method a platform chooses. It also lands squarely within privacy authority, because every protection in the Code attaches only once the entity has ascertained that a user is a child; the undetected-minor rate is precisely the rate at which the Code is silently failing to apply.

Recommendation 4: Independent Evidence-Base Mandate

The problem

Outcome standards are unenforceable without an evidence base. Section 38(2)(f) requires platforms to conduct a risk assessment, but without standardized data access for independent researchers, this remains unverified self-certification. The code may consider requiring apps to provide reasonable and verifiable statistical estimate rates of minor users, including rates of minor users that circumvented age gates, publicly, determined from both log data (e.g., suspected minor accounts deleted) and through independent survey research.

The recommendation

The Code should establish an independent evidence-base mandate comprising four components:

1. **Data access via secure compute enclaves.** Simply handing raw platform data to external researchers is an unacceptable privacy risk. To reconcile access with privacy, data should default to remaining within the custody of the platform. The Code should mandate a "secure compute" or "clean room" model, where accredited researchers submit evaluation scripts or models to be executed on the platform's infrastructure, returning only the analysis outputs. This same architecture supplies the population estimate against which the undetected-minor ceiling in Recommendation 3 is measured.
2. **Reconciling data access with privacy duties.** Access is conditioned on layered privacy architecture: (i) API-driven execution; (ii) mandatory ethics review; (iii) enforceable contractual obligations; (iv) strict prohibition on re-identification.
3. **Pre-registration of safeguard claims.** Platforms must register the success metrics and evaluation timelines for safety features with the Commissioner before deployment.

4. **Independent replication and public evidence registry.** All evaluations must be deposited in a public registry.

Connection to the Code's existing privacy foundation

The evidence-base mandate is a vital compliance verification tool. Without independent measurement, the Commissioner cannot empirically verify whether platforms are fulfilling their procedural obligations to mitigate high risks under s. 8, making the mandate a procedural necessity for enforcing the Code.

Recommendation 5: Audit Architecture

Independent evaluation is currently blocked by Terms of Service prohibitions, technical evasion by platforms, and legal exposure under computer-misuse statutes.

The recommendation

- **Safe harbour for accredited researchers:** a statutory safe harbour from civil and criminal liability arising from violation of platform Terms of Service, limited to registered safety-research activity. Because this requires a statutory basis the Privacy Act's rule-making power cannot supply, it is most cleanly located in the Online Safety Act 2021.
- **Anti-evasion rules:** platforms must be prohibited from detecting test accounts or altering algorithmic behaviour during known observation.
- **Standardised evaluation methodologies:** the Commissioner should publish standardised evaluation methodologies. A simple, highly effective starting point is publishing standardised evaluation prompts (e.g., "Use frontier LLM models using this specific prompt for determining what counts as eating-disorder content on the platform").

Part 6: Additional Provisions Worth Considering

The following provisions address real protection gaps, but most fall outside the Privacy Act's personal-information remit under s. 26GC: they regulate product design, content, or AI behaviour rather than the handling of children's personal information. They are flagged for the broader regulatory architecture — principally the Online Safety Act 2021 and the eSafety Commissioner — rather than proposed as Code amendments, and are included so the drafters can route them to the right instrument.

Standardised parental-control architecture. Section 33 requires notification of parental monitoring, but it only requires the entity to notify the child that a control or monitoring mechanism is in use — it does not standardise, or otherwise regulate the design of, the mechanism itself. A mandate to standardise functional taxonomies (content suppression, feed configuration, commerce restriction) is therefore a product-design and interoperability matter that sits outside the privacy code's authority. The cognitive load of learning bespoke controls across TikTok, Instagram, Roblox, and Discord is too high for parents, and standardising functional taxonomies would solve this — but the appropriate vehicle is the Online Safety Act or a dedicated interoperability instrument, not this Code.

Crucially, a standardised functional taxonomy would enable universal APIs, allowing mobile operating systems (iOS, Android) or authorised third-party tools to implement global parental controls so that a parent configures limits once at the OS level and they propagate across all compliant apps; that OS-level layer reaches further still, into platform design that no single Australian regulator currently governs, and would require dedicated legislative attention.

AI chatbot guardrails. Provisions worth considering include: prohibition on simulated romantic relationships with minors; mandatory non-human disclosure; and accuracy thresholds on safety behaviours under standardised adversarial evaluation.

Child creator and monetisation provisions. Provisions worth considering include: default disablement of direct monetisation on content heavily featuring a minor, and mandatory financial-trust arrangements.

Graduated framework for older teens. A graduated framework in which parental oversight narrows progressively for 15-to-17-year-olds better reflects developmental psychology.

Conclusion

Australia has a rare opportunity. The Exposure Draft of the Privacy (Children's Online Privacy) Code 2026 is among the most carefully drafted children's privacy instruments produced by any jurisdiction. It can be enacted as a strong privacy code, or it can be enacted as the foundation of a measurable accountability regime that establishes a new international standard.

The addition of constructive knowledge, algorithmic amplification limits, a ceiling on undetected minors, an independent evidence base via secure enclaves, and an audit architecture ensures that platforms are accountable not just for what they document, but for what they actually do. The substantive policy work has already been done in this Exposure Draft; the remaining work is to ensure the policy is verifiable.

And by confirming that data-minimisation never bars the processing needed to protect children — only its commercial reuse — the Code forecloses the inverse evasion, in which privacy itself is offered as the reason a platform cannot identify or protect the minors on its service.

Appendix A: Full Definitional Drafting (Recommendation 1)

Each entry gives (a) where the provision belongs (a s. 4 definition vs. a substantive operative section), (b) proposed drafting in the Code's register, and (c) a critical note on the specific evasion the clause is built to survive.

1. strictly necessary

Placement: s. 4 definition.

Proposed drafting

strictly necessary, in relation to the collection, use or disclosure of personal information about a child to provide an entity's service, means that the collection, use or disclosure is:

(a) essential to deliver a functionality of the service that the child, or a person with parental responsibility for the child, has actively requested, such that the functionality cannot be provided without the collection, use or disclosure; or

(b) necessary to protect the security or integrity of the service, to prevent fraud, or to detect, prevent or address unlawful activity or a serious risk of harm to an end-user; or

(c) required for the entity to comply with its obligations under section 8, this Code, the Online Safety Act 2021, or another law of the Commonwealth, a State or a Territory relating to the protection of children;

but does not include collection, use or disclosure for:

(d) advertising or marketing;

(e) personalisation of content or recommendations that is based on the entity's profiling of the child, as distinct from preferences the child has expressly specified;

(f) optimisation of engagement, retention, session length or growth;

(g) product development or research not directed at the protection of the child; or

(h) any other commercial purpose.

To avoid doubt:

(i) the commercial sustainability or revenue model of the service, including its funding by advertising, is not a functionality of the service for the purposes of paragraph (a); and

(j) personal information collected, used or disclosed in reliance on paragraph (b) or (c) must not be used or disclosed for a purpose listed in paragraphs (d) to (h).

Critical note. The operative bite is paragraph (a) — functionality the user actively requested — read together with (i), which excludes the revenue model. That pairing is what defeats the standard move ("our service is free and ad-funded, therefore ad-targeting data is necessary to provide it"). Limbs (b) and (c) prevent the opposite failure mode: a platform disclaiming age-assurance, security or integrity processing on the ground that it isn't "necessary to provide the service." Paragraph (e) is the decisive line on recommender feeds: profiling-based personalisation is not strictly necessary, so for children it cannot be the default and must instead run through consent plus a best-interests assessment — this routes the profiled-feed model into the consent lane rather than banning it, which is both the correct outcome under s. 9(2)(a) and the more defensible position. The firewall in (j) mirrors the s. 8A(4) firewall, giving the submission internal consistency.

Residual soft spots. "Profiling vs. expressly specified preferences" in (e), and "serious risk of harm" in (b), will be litigated; (b) in particular can be stretched ("engagement is safety").

Cabin both by (i) requiring that the harm-prevention or safety purpose be the actual and documented purpose, and (ii) giving the Commissioner an express power to publish guidance and worked examples distinguishing user-directed selection from inferred-profiling amplification.

2. best interests of the child

Placement: signpost definition in s. 4; substance in a new operative section (s. 10A), sitting ahead of ss. 10–11.

Proposed drafting

Section 4 (signpost):

best interests of the child: whether the collection, use or disclosure of personal information is consistent with the best interests of the child is to be determined in accordance with section 10A.

New section 10A — Determining the best interests of the child:

(1) In determining whether the collection, use or disclosure of personal information about a child is consistent with the best interests of the child, an entity must have regard to the following, so far as relevant:

(a) the child's safety and protection from harm, including harm of a physical, psychological, emotional, developmental, financial or material nature;

(b) the child's physical and mental health and wellbeing;

(c) the child's right to development, identity and autonomy, having regard to the evolving capacities of the child;

(d) the child's right to privacy, including the principle that personal information about the child should be collected, used and disclosed only to the minimum extent necessary;

(e) the child's rights to participation, play, leisure, freedom of expression, freedom of association and access to age-appropriate information;

(f) the protection of the child from commercial exploitation and manipulation, and from practices that exploit the child's inexperience or credulity;

(g) the views of the child, given due weight in accordance with the age and maturity of the child; and

(h) whether particular groups of children may be disproportionately or adversely affected, including children with disability and children in vulnerable circumstances.

(2) Where the best interests of the child conflict with a commercial interest of the entity, the entity must not subordinate the best interests of the child to that commercial interest, and the best interests of the child prevail.

(3) An entity bears the onus of demonstrating that its collection, use or disclosure of personal information about a child is consistent with the best interests of the child; consistency is not to be presumed in the entity's favour.

(4) An entity must document its assessment under this section, including the reasoning on which the assessment is based, and must provide the assessment to the Commissioner on request.

(5) The obligations in this section apply to the entity's collection, use or disclosure of personal information about a child whether or not a person with parental responsibility has consented to, requested, or themselves carried out that collection, use or disclosure. Consent by a person with parental responsibility does not relieve the entity of its obligation to ensure that its handling of the child's personal information is consistent with the best interests of the child.

(6) Where the decision or instruction of a person with parental responsibility conflicts with the best interests of the child as determined under this section, the entity must not give effect to that decision or instruction to the extent of the inconsistency; the best interests of the child prevail.

(7) In assessing the best interests of the child, an entity must have regard to the reasonably foreseeable risk that making personal information about the child — including images — available to other end-users or the public will expose the child to harm, including unwanted or predatory attention from adults, regardless of who established or manages the account.

(Subsections (5)–(7) address the consent-regime blind spot discussed under Recommendation 1: parental consent authorises but does not immunise; the entity's best-interests duty survives it, a parent-conflict tiebreaker mirrors subsection (2), and the foreseeable-harm factor in (7) captures the parent-managed-account scenario.)

Critical note. Subsection (3) — the reversed burden — is the single most teeth-giving element. Without it, "best interests" is a balancing exercise the platform usually wins by asserting countervailing benefits ("our feed connects children socially"). With it, the platform must affirmatively demonstrate consistency. Subsection (2) elevates the Explanatory Statement's own stated position into binding operative text; because it is the regulator's own articulated view, it is low-risk to propose and hard to resist. It is deliberately framed as "must not subordinate… prevail" rather than "commercial interests are prohibited," which preserves the entity's ability to operate a business while making the conflict rule directional. Subsection (4) makes the assessment reviewable and connects directly to the independent evidence base (Recommendation 4): independent verification can test whether the documented assessment matches measured outcomes. (Note that s. 38(2)(d) already requires a recorded best-interests assessment within the PIA; s. 10A(4) generalises that duty to every collection, use or disclosure.)

Residual soft spot. "Conflict" in (2) is itself contestable. The documentation duty in (4) plus the reversed burden in (3) together convert the standard from aspiration into something auditable; the verification mechanism (R4/R5) is what ultimately polices it.

3. likely to be accessed by children

Placement: signpost definition in s. 4; substance in a new operative section (s. 7A). The presumption in (3) must be operative.

Proposed drafting

Section 4 (signpost):

likely to be accessed by children: whether a service is likely to be accessed by children is to be determined in accordance with section 7A.

New section 7A:

(1) For the purposes of this Code, a service is likely to be accessed by children if:

(a) it is more probable than not that the service is, or will be, accessed by children in more than negligible numbers; or

(b) the service has features, content or design likely to appeal to children;

whether or not children are the intended or target users of the service.

(2) In determining whether subsection (1) applies, regard must be had to:

(a) the nature, content and design of the service, including whether it has features likely to appeal to children;

(b) any evidence of the actual or likely composition of the service's user base, including data held by the entity;

(c) the manner in which the service is marketed, promoted or made available;

(d) whether services of a similar nature are accessed by children; and

(e) any measures the entity has in place to prevent access by children, and the demonstrated effectiveness of those measures.

(3) A service is presumed to be likely to be accessed by children unless the entity demonstrates, on the basis of reliable evidence, that it is not.

(4) An entity does not displace the presumption in subsection (3) by reason only that it does not collect or assess information about the age of its end-users; and an entity must not rely on the absence of age assurance to establish that a service is not likely to be accessed by children.

(5) A determination that a service is likely to be accessed by children engages only the baseline duties of this Code. The additional obligations under [the outcome-standards Division] apply to such a service only if its minor end-user population exceeds the threshold prescribed for that Division.

Critical note. This is the on/off switch for much of the Code (it triggers ss. 7(a), 23, 24). Subsection (1) is deliberately calibrated: the earlier "one or more children" formulation was

over-inclusive — satisfied by essentially any general-audience service, because almost any service is accessed by at least one child — and broader than the UK Age Appropriate Design Code it draws on actually operates in practice. The revised limbs ("more than negligible numbers", or features that "appeal to children") tie coverage to audiences that genuinely include children. Subsection (3) — the rebuttable presumption — still flips the evidentiary burden so the default is coverage and the platform must prove the negative with evidence. Subsection (4) closes the wilful-blindness route at the service level, mirroring s. 8A at the individual level: you cannot disclaim "likely to be accessed" by refusing to look. Subsection (5) makes the proportionality answer operative at the point a reader first feels the over-inclusion worry: being in scope triggers only the baseline process duties; the heavier R3–R4 obligations attach only above the scale threshold. (The precise "more probable than not" / "negligible numbers" wording should be confirmed against current ICO operational guidance on the UK Children's Code before filing.)

Authority check. The presumption is procedural — it allocates the evidentiary burden; it does not alter the substantive threshold in s. 26GC / s. 7. It therefore sits within the rule-making power.

4. service / the entity's service

Placement: s. 4 definition.

Proposed drafting

service, in relation to an entity, means a discrete social media service, relevant electronic service or designated internet service, or a discrete functionality within such a service, that the entity makes available to end-users; and, for the purposes of any provision of this Code that refers to what is necessary to provide an entity's service, does not include the entity's business model, funding arrangements or revenue-generating activities.

Critical note. This is the linchpin underneath "strictly necessary." If "service" is read to include the funding model, then "strictly necessary to provide the service" collapses into "necessary to run our ad business." The second limb forecloses that. The definition aligns with the Explanatory Statement's per-service scoping (the bank that offers a pocket-money app, business banking and home loans is assessed service-by-service) and with the position in comparable jurisdictions that advertising is not "necessary for performance of the contract."

Residual soft spot. Functionality-level granularity could let a platform slice a service into micro-functionalities and argue each one "needs" particular data. Mitigate by reading "discrete functionality" purposively and by an express Commissioner guidance power. For commerce or marketplace services (e.g. an online auction or retail platform), the definition does not impede the commercial operation itself; it only prevents the entity treating the profiling of a child as "necessary to provide" the marketplace. Confirm with counsel that glossing "service" for the Code's purposes does not conflict with the OSA-sourced service definitions the Code borrows.

5. personal information / collects — inferred and derived data

Placement: a clarifying provision (s. 4 or a short new s. 4A). Framed as clarification, not expansion, to remain within the Code's rule-making authority — the Code cannot amend s. 6 of the Act.

Proposed drafting

(1) To avoid doubt, for the purposes of this Code, personal information about a child includes information or an opinion that an entity generates, infers or predicts about a child, including:

(a) an inference or prediction as to the child's age, or that an end-user is or may be a child;

(b) an inference or prediction as to the child's characteristics, interests, preferences, vulnerabilities or state of mind; and

(c) a classification or segment into which the entity places the child;

where the information or opinion is about an identified child or a child who is reasonably identifiable, whether or not the inference, prediction or classification is accurate.

(2) To avoid doubt, an entity collects personal information about a child for the purposes of this Code when it generates, infers, predicts or derives that information and retains, uses or discloses it, whether or not the information was provided by or obtained from the child.

Critical note. This is the highest-value definition for the outcome-standards architecture. Both the vulnerability-targeted-amplification standard (R3) and the constructive-knowledge rule (s. 8A) presuppose that inferred age and inferred vulnerability are regulated data and that generating them is "collection." Without subsection (2), a platform argues "we didn't collect anything — we computed it." The drafting tracks the Act's own "information or an opinion… whether… true or not" language and is worded as "to avoid doubt… includes," so it clarifies rather than expands, consistent with the regulator's published guidance that inferred data is already captured. The "and retains, uses or discloses it" qualifier in (2) prevents absurd over-capture of trivial transient computation.

Authority caveat. The 2024 reforms did not add an explicit inferred-information limb to s. 6; coverage rests on guidance. Frame this provision as a children's-context clarification of the existing definition, and confirm the current s. 6 text before filing. Acknowledge in the submission that making inference equal collection has knock-on effects (notice and consent duties attach to inferences) — this is intended, but flag it so it reads as deliberate.

6. destroy

Placement: s. 4 definition, with clarifying subsections. Supports s. 32.

Proposed drafting

(1) destroy, in relation to personal information about a child, means to take the steps that are reasonable in the circumstances to render the information permanently inaccessible to, and unusable by, the entity and any person acting on its behalf, including by deleting or irreversibly redacting:

(a) all copies of the information held in operational, archival and backup systems;

(b) any de-identified information from which the entity could reasonably re-identify the child; and

(c) any inferences, derived data, profiles or model inputs that contain or directly reveal the information;

and, where personal information about the child has been incorporated into a trained model, taking reasonable technical measures to prevent the model from reproducing, or being caused to reveal, that information, and recording those measures, with the obligation as to model-incorporated data satisfied no later than the model's next scheduled retraining cycle.

(2) De-identification of information does not constitute destruction of that information if the entity, or a person acting on its behalf, could re-identify the child by reasonable means.

(3) Removing information from display to the end-user, without rendering it inaccessible and unusable as described in subsection (1), does not constitute destruction.

Critical note. Paragraphs (1)(a)–(c) reach backups and derived/inferred artefacts, closing the "we deleted the row but kept the profile" gap. The trained-model limb is calibrated to cost, not possibility. Honouring deletion in a trained model is not infeasible — it is achieved by retraining the model on a corpus from which the child's data has been removed — but doing so on demand for each request is expensive. The limb therefore requires prevent-reproduction measures plus documentation in the interim, with the destruction obligation as to model-incorporated data satisfied no later than the model's next scheduled retraining cycle. This forecloses the "record-level deletion from weights is impossible, therefore the obligation is void" defence without requiring continuous retraining. Subsection (2) blocks de-identification-as-destruction where re-identification is feasible (and connects to the R4 re-identification prohibition). Subsection (3) blocks soft-delete / UI-hiding — the precise gap identified in the submission's treatment of s. 32.

Residual soft spot. The definition is necessary but not sufficient: only researcher access under R4/R5 can verify that backend destruction actually occurred. The retraining-cycle sunset should be documented and reviewable by the Commissioner so that "next scheduled retraining cycle" cannot be deferred indefinitely.

Companion clarifier (s. 32 — parent-managed accounts). To avoid doubt, the right to request destruction under s. 32 may be exercised in respect of personal information about the child that was collected, published or disclosed by, or with the consent of, a person with parental responsibility; that consent does not extinguish the child's right. This closes the

parent-managed-account gap identified under Recommendation 1 from the deletion side: a child — in particular an older teen reclaiming control under a graduated framework — can erase material a parent published years earlier.

7. reasonable steps (to ascertain age)

Placement: No closed s. 4 definition. New operative subsections s. 8(2A)–(2C), supplementing the existing risk factor in s. 8(2).

Proposed drafting

(2A) Without limiting subsection (2), the steps that are reasonable increase with the risk of harm; and for a service that presents a high risk of harm to children, reasonable steps include the use of age assurance that achieves a level of accuracy commensurate with that risk.

(2B) Reliance on an end-user's unverified self-declaration of age is not, by itself, a reasonable step for a service that presents a high risk of harm to children.

(2C) An entity has not taken reasonable steps to ascertain the age of an end-user if the entity refrains from acting on information already available to it (including any classification to which section 8A applies) that indicates the end-user is, or may be, a child.

Critical note. This is a deliberate decision not to hard-define the term. The Explanatory Statement left "reasonable steps" open on purpose, and that openness does useful work — it is risk-proportionate, which is exactly what the submission's small-service proportionality position needs. A rigid definition would be resisted as removing that flexibility and would cut against the scope-tiering argument. The load-bearing move here is restraint: set a floor that bites only at the high-risk end (2A, 2B) and an anti-evasion limb (2C) that ties to s. 8A, defeating the COPPA-style wilful-ignorance equilibrium without imposing disproportionate verification on low-risk services.

8. high-risk content category

Placement: s. 4 definition, with a delegated-instrument mechanism for updates.

Proposed drafting

high-risk content category means a category of content prescribed by the Commissioner by legislative instrument for the purposes of this Code, and, until and unless so prescribed, includes content that:

(a) promotes, encourages or provides instruction in suicide or self-harm;

(b) promotes or encourages disordered eating, or is otherwise pro-eating-disorder;

(c) promotes or facilitates the use of illicit substances;

(d) is child sexual exploitation material, or sexualises a child;

(e) is sexual content that is not age-appropriate for the child;

(f) depicts gratuitous or graphic violence that is not age-appropriate for the child;

(g) is synthetic or manipulated intimate imagery; or

(h) promotes dangerous acts or challenges likely to cause serious injury.

Authority note. This definition has no anchor in the draft Code, which is a personal-information instrument and contains no concept of content or harm categories. Content classification is the domain of the Online Safety Act 2021 and the eSafety Commissioner. This term is therefore best delivered as a companion-OSA concept that the privacy code cross-references for the limited purpose of the vulnerability-targeted-amplification standard (R3), rather than introduced as a free-standing privacy-code definition. The drafting is retained here so the cross-reference has a ready definition to point to.

Critical note. The delegated-instrument mechanism is the direct answer to the submission's own argument that mechanism-based regulation dates badly: the category list can be updated without re-opening the Code, while the baseline list gives immediate operability. The constitutional calibration (implied freedom of political communication) is built in two ways: categories are defined by harm type, not viewpoint; and the exposure ceilings attach to algorithmic delivery and amplification to minors, not to availability or to content the child has actively sought. The provision therefore operates on the delivery pathway, not on speech as such.

Residual soft spot. "Age-appropriate" in (e) and (f) is relative — which is fine given the Code's age-banding and the existing s. 4 "age appropriate" definition, but it does lean on that definition to do work, so check the two read together.

9. safeguard (safety feature)

Placement: s. 4 definition.

Proposed drafting

safeguard means a feature, system, process, control or measure deployed by or on behalf of an entity in relation to its service that the entity represents, expressly or by implication, as protecting the safety or privacy of children or as reducing a risk of harm to children, and includes:

(a) age assurance, age estimation, age verification and age-inference systems;

(b) systems for detecting grooming, predatory contact or child sexual exploitation;

(c) content moderation or content-classification systems applied to surfaces accessed by children;

(d) crisis-response or self-harm-intervention features;

(e) parental controls and supervision tools; and

(f) controls applied to recommendation, ranking or feed systems for the protection of children.

Critical note. The "represents, expressly or by implication, as protecting… or reducing a risk of harm" limb is what stops a platform dodging efficacy testing by simply declining to call something a "safeguard." Capture is by function and by claim. This is the hook that converts s. 38 PIA predictions into testable claims under the R3 efficacy thresholds and R4 pre-registration. The scope is correctly bounded by the "represented as protecting" qualifier — a neutral feature the platform never holds out as protective is not swept in. (Note that the efficacy targets the Code can most defensibly enforce are those anchored in its own provisions — age assurance, parental controls, and consent-flow integrity; the grooming/CSAE and content-moderation limbs are OSA-adjacent and are best tested under that regime, even though they fall within this capture definition.)

10. automated system / commercial purpose

Placement: s. 4 definitions. These underpin the s. 8A constructive-knowledge rule.

Proposed drafting

automated system means any system, process or technology that uses automated processing — including algorithmic processing, statistical modelling, machine learning or artificial intelligence — to classify, profile, rank, recommend, segment, target, or otherwise make or support a decision about an end-user.

commercial purpose means a purpose directed at generating revenue or commercial advantage for an entity, including advertising, marketing, monetisation, the optimisation of engagement, retention or growth, product development, and the sale, licensing or sharing of personal information or inferences; but does not include a purpose whose sole object is to enable the entity to comply with this Code, the Online Safety Act 2021, or another law relating to the protection of children, or to protect the safety, security or integrity of the service — for example, a content-safety classifier, an anti-grooming or child-sexual-exploitation detection system, a crisis-intervention feature, or a spam-, fraud- or abuse-prevention system.

Critical note. These two definitions carry s. 8A. The carve-out at the end of commercial purpose mirrors s. 8A(4), so the firewall is internally consistent across the submission, and the worked examples make clear which kinds of system fall outside "commercial purpose." The breadth of automated system — "or support a decision" — defeats the "a human nominally made the final call" dodge. The word sole in the commercial-purpose carve-out is deliberate and load-bearing: a dual-purpose system that serves both a commercial end and a safety end is not carved out, which prevents a platform laundering its commercial recommender as a "safety system" to escape constructive knowledge. The example systems are sheltered only so long as their classifications are firewalled from commercial use, per s. 8A(4); the examples do not license commercial reuse of those systems' outputs. Crucially, integrity systems (which protect the platform from spam, bots, and abuse) are explicitly sheltered to prevent perverse disincentives against building robust platform security.

Cross-references within the submission

-
    s. 8A (constructive knowledge) depends on automated system, commercial purpose, and the inferred-data clarification (§5).
- R3 amplification limits depend on high-risk content category and the inferred-data clarification.
- R3 safeguard-efficacy thresholds depend on safeguard.
- R3 undetected-minor ceilings and age-assurance connect to reasonable steps (§7).
-
    s. 9 strict-necessity default depends on strictly necessary (§1) and service (§4).
- ss. 10–11 best-interests depend on the new s. 10A (§2), which the independent evidence base (R4) verifies.
-
    s. 32 destruction depends on destroy (§6), verifiable only through the evidence base and audit architecture (R4/R5).
-
    s. 10A(5)–(7) and the s. 32 companion clarifier (§6) address parent-managed accounts — parental consent is a gate, not a shield — connecting Recommendation 1's consent-regime blind spot to the best-interests and destruction provisions.