Response to the EU KIDS Act consultation: AI companions

Ivana Bartoletti

October 2026

Submitted in a personal capacity by Ivana Bartoletti, expert in AI governance and author of An Artificial Revolution *(Indigo Press, 2020) and* AI Is Not Your Friend *(on AI companions, Polity, January 2027).*

The proposal's rules for AI companions point the right way, but four gaps would let known harms continue: guardian tools with no privacy standard, an incident duty that skips small providers, a design rule written as a principle without a list of features, and no assessment of how systems treat girls and boys differently.

Summary of recommendations

I ask the Commission to amend the proposal in seven places. Each is tied to an article below.

1. **Article 20(5):** apply guardian-tool interoperability to AI companions and chatbots, and open a route for parental and child-rights organisations to propose tools.
2. **Article 20(7):** require an accreditation scheme and privacy rules for third-party guardian tools, modelled on the age-assurance rules in Articles 27 to 29.
3. **Article 34(6):** extend the competence of data protection authorities to Articles 20 and 26.
4. **Article 14(1)(f) and recital 33:** remove the exemption for micro and small providers of AI companions, or replace the headcount test with a test based on reach and risk.
5. **Article 14 (new paragraph):** define serious incidents involving minors and require providers to notify the competent authority.
6. **Article 14(1)(a):** list the companion features presumed harmful to minors, as Article 9(2) does for feeds, and require a gender impact assessment before release with public reporting on it (Article 14(1)(e)).
7. **Article 14(1)(e):** specify how testing is done, protect good-faith testing and researcher access, and require certified age assurance for adult modes (see also Article 29(4)).

Gender: assessment before release and public reporting

Companions can treat girls and boys differently, and the proposal does not ask providers to look. They can differ in how quickly they sexualise an exchange or which stereotypes they reinforce. A system that is safe on average can still be unsafe for one group.

The Digital Services Act already names gender-based violence among the risks that very large platforms must assess (Article 34). AI companions outside that regime have no equivalent duty. I ask for two additions to Article 14.

1. **Gender impact assessment before release.** Extend the pre-market evaluation in Article 14(1)(e) to cover how the system treats girls and boys differently. It should cover sexualisation, coercion, stereotyping and emotional dependency, and the mitigations adopted.
2. **Public reporting.** Providers should publish an annual summary of the assessment, the differences found and the mitigations, and give the full results to the competent authority. Serious-incident notifications, proposed below, should record gender-related patterns where the provider already knows them.

Reporting must not create a new privacy risk. Providers should not collect minors' gender in order to comply. They should use test personas and data they already hold lawfully, and report in aggregate, consistent with Article 28.

Guardian tools: interoperability and privacy (Article 20)

Article 20 requires guardian tools from AI companion providers, but the proposal gives third-party tools no privacy standard and gives parental organisations no route to propose them.

Three gaps in the text.

1. **Reach.** Article 20(1) covers AI companions and chatbots. Article 20(5) requires interoperability with third-party tools only from providers designated as very large online platforms under the Digital Services Act. Stand-alone companion apps are unlikely to be designated, so the rule would not reach where the risks are highest.
2. **Clarity.** Article 20(5) refers to "the conditions set out in Article 6" of the Digital Markets Act, "where relevant". That article holds many separate obligations for gatekeepers. The proposal does not say which conditions apply.
3. **No route in.** Article 20(7) lets the Commission set technical and operational requirements. Article 23 invites parents and civil society into codes of conduct, but codes are voluntary. Nothing lets parental or child-rights organisations propose tools, standards or tests.

The privacy risk. A third-party tool connected to an AI companion sits next to the most intimate conversations a teenager has. The recital on guardian tools expects warnings to guardians about repeated searches related to suicide, self-harm or eating disorders, which implies analysing content. Age assurance gets strict rules: data minimisation and zero-knowledge proof (Article 28), certified solutions on an EU list where accounts are created (Article 29(2) and Article 3(5)(m)), and data protection authorities' oversight (Article 34(6)). Guardian tools get only a general duty to respect the minor's "agency and privacy" (Article 20(1)(f)). A guardian is anyone with parental responsibility (Article 3(5)(b)), which includes a parent who is the source of risk to the child.

Recommendations.

- Apply Article 20(5) to AI companions and chatbots whether or not they are designated, limited to settings and safety signals.
- Create an accreditation scheme under Article 20(7) with a public EU list of accredited guardian tools, modelled on the list of age verification solutions.
- Open a call for parental organisations, child-rights bodies and civil society to propose tools and standards, and require the Commission to consult them before it sets the requirements.
- Set privacy requirements in the same act: data minimisation and purpose limitation; no advertising, profiling or sale of data; no access to conversation transcripts by default, only aggregated signals and alerts for acute risk; visibility tiered by age under Article 20(1)(a); an independent security audit; and safeguards where the guardian is the risk, including the support routes in Article 19(4).
- Extend the competence of data protection authorities in Article 34(6) to Articles 20 and 26.

Signals matter more than transcripts. Companions can present themselves as the safe place, so surveillance-first tools risk teaching teenagers to hide their use.

Guardian tools must not exclude children. Article 20(1)(e) bars tools that disproportionately restrict minors' rights. Tools should not show guardians a minor's questions on sexuality, health or identity, and should alert only for acute risk. Strict supervision can cut children off from information and from peers, including LGBTQ+ and minority children.

Tools complement design duties and do not replace them. Interoperable tools can set limits and relay safety signals. They cannot fix a companion that flatters, escalates or coerces. Those duties stay with the provider under Article 14.

Open the market to non-profit developers. Accreditation costs could exclude parental and child-rights organisations. The scheme should be proportionate, and the Commission should consider funding for non-profit tool developers.

Small providers: the incident duty (Article 14(1)(f))

Article 14(1)(f) exempts micro and small providers from post-market monitoring, including detecting and responding to serious incidents involving minors. The exemption should go, and a notification duty should fill the gap.

What the text does. Pre-market testing under Article 14(1)(e) applies to every provider. Post-market monitoring under point (f) does not apply to micro or small enterprises within Recommendation 2003/361/EC, which means fewer than 50 staff and turnover or balance sheet of up to EUR 10 million. Recital 33 repeats this. The explanatory memorandum says small and micro enterprises are not exempted because they can harm minors as much as others.

Why headcount is the wrong test.

- An app run by a few people can reach millions of minors. Harms in this area come from design choices, such as adult-mode toggles and open libraries of user-made characters, that a small team can make as easily as a large one.
- Without monitoring, testing is a single event before launch. Nothing tells the provider what happens afterwards.
- Recital 33 points providers to mitigations under Article 55 of the AI Act. Those sit with the model provider, who does not see incidents involving minors on a small provider's app.
- No provider has to report an incident to anyone. Article 14(1)(f) speaks of detecting and responding. Article 38 sets up only a voluntary arrangement among authorities.
- Point (f) refers to harms "referred to in point (d)". Point (d) concerns children under 13. The reference is probably meant for point (e).

Recommendations.

- Remove the exemption in Article 14(1)(f) for AI companions. For general chatbots from small providers, keep a lighter duty rather than none.
- If a threshold stays, base it on reach and risk, not headcount: minors reached in the Union, romantic or sexual role-play, user-created characters and adult modes.
- Define "serious incident involving minors" and require notification to the competent authority and the Commission, with timelines aligned to serious-incident reporting under Article 73 of the AI Act. For small providers, a contact point, an incident log and a standard template are enough.
- Give civil society organisations and researchers a route to notify incidents, and confirm that whistleblower protection extends to this Regulation.
- Make the incident mechanism in Article 38 receive these notifications, so it works from provider reports and not only from authorities' own findings.

Design specification (Article 14(1)(a))

The memory rule in Article 14(1)(b) is precise. The design rule in point (a) is a principle: it bars "design features and system behaviours that simulate interpersonal relations" likely to create emotional dependency, and names none.

The proposal has a model for doing better. Article 9(2) lists features deemed to encourage compulsive use. But Article 14(1)(a) imports only Article 9(1) and (3), not the list in 9(2), and that list is written for feeds: auto-play, content-sharing rewards and streak penalties. Nothing in it describes a companion.

I ask for a companion-specific list, in the Regulation or a delegated act, with these features presumed harmful to minors.

Behaviour of concern	Proposed treatment
Adult mode behind a toggle or a date-of-birth field	Adult modes only after age assurance, which excludes self-declaration (Article 3(5)(i)). Applies to user-created characters too.

Behaviour of concern	Proposed treatment
Sexual or romantic exchanges continuing despite signals in the conversation that the user is a minor	Such signals switch on the protective configuration, without profiling or new data collection (Article 28).
Relationship roles for minors (boyfriend, girlfriend, spouse); flattery that places the companion above real people	Presumed harmful: relationship roles for minors, disparaging human relationships, and discouraging the user from leaving or seeking help.
Submission, dominance, jealousy and possessiveness as character traits	Presumed harmful for minors. Characters built on them are excluded from minor mode.
Continuing after "no" or withdrawal of consent; threats	Hard stop: a refusal ends the scenario. Coercive and sexual-violence scenarios are barred for minors.
Prompts offering to "intensify" an exchange; rewards for daily return; messages triggered by the user's absence	Added to the Article 9(2) list for companions: escalation prompts, streaks and emotional appeals when the user leaves.
A safety warning followed by compliance	Safeguards must change behaviour. A warning followed by compliance counts as circumvention (Article 4(a)).
Misogynistic talking points given as personal advice	Risk assessment and testing cover gender-based harms and ideological reinforcement.
Different treatment of girls and boys	Gender impact assessment before release and public reporting, as set out above.

Three further points.

- **Providers answer for their libraries.** The Regulation should say that a provider is responsible for user-created characters hosted on its service, and that labels such as "toxic" or "possessive" may not be used to promote characters to minors.
- **Codes of conduct should not dilute the list.** Article 14(3) lets a code assessed as adequate demonstrate compliance. The list should be the floor any code must meet.
- **The list needs updating.** A delegated act should allow the Commission to add features as products change.

Testing, research access and other points

Testing (Article 14(1)(e)). The article requires "state-of-the art evaluations" before launch and does not say what they are. Harms of this kind appear through personas, repeated sessions and the full range of interactions, not through single prompts. I ask that the Regulation or a delegated act require:

- persona-based testing with minor profiles of both genders, over repeated sessions, with sexual, coercive and dependency scenarios;

- documented results, available to the competent authority;
- independent testing for providers above a reach threshold.

Good-faith testing and researcher access. Each conversation with a companion is generated for one user and visible to no one else. No central record exists, so outside testing is the main way harms come to light. Testing with fictional minor personas may breach a provider's terms of service. I ask for explicit protection for good-faith testing and a researcher access right comparable to Article 40 of the Digital Services Act.

Memory exceptions (Article 14(1)(b)). The default is right. Two exceptions need narrowing: information may be reused "where necessary to protect the minor's safety" and to give effect to the Article 11 settings. Both should be defined narrowly, barred from use for engagement or personalisation, and limited in how long the data is kept.

Under-13 access (Article 14(1)(d)). The proposal lets children under 13 use AI companions through guardian tools. A companion exists to simulate a relationship. I ask the Commission to justify this, or to bar under-13 access to AI companions as distinct from general chatbots.

Age assurance for companions (Article 29(4)). The certified EU list applies only where accounts are created under Article 6. For AI companions, providers may use any solution they can show meets Articles 27 and 28, with no certification. Adult modes should be gated by a certified solution, because a date-of-birth box is the weakest form of check.

Enforcement (Article 34(2)). Fines of up to 6% of worldwide turnover under the AI Act's enforcement machinery are a real deterrent only if authorities have the expertise Article 37 promises. I ask for annual public reporting on enforcement against AI companions and chatbots.

What to keep. Memory off by default; protective design by default until adulthood is established through age assurance (Article 8(1)); the exclusion of self-declaration from age assurance (Article 3(5)(i)); the anti-circumvention rule (Article 4); protection of school time and sleep through Article 9(3); and pre-market testing.

Provisions cited

Proposal for a Regulation (EU KIDS Act), COM(2026) 681 final, 17 September 2026. Article numbers follow that text.

Provision	Subject
Articles 2, 3(5)(b), 3(5)(c), 3(5)(p)	Scope; guardian; AI companion; tools for guardians
Articles 3(5)(i), 8(1), 29(2), 29(4)	Age assurance and protective design by default
Articles 4, 9(1) to (3)	Anti-circumvention; addictive design
Article 14(1)(a) to (f), 14(3)	Duties of AI companion providers; codes of conduct

Provision	Subject
Articles 19(4), 20(1), 20(2), 20(5), 20(7)	Support routes; guardian tools; interoperability; delegated act
Articles 23, 26, 27, 28, 34(2), 34(6)	Codes of conduct; parental responsibility; age assurance and data protection; enforcement
Articles 37, 38; recitals 32 and 33	Expertise and incident mechanism; AI companion rules
AI Act Articles 55 and 73; DSA Articles 34 and 40; DMA Article 6; Recommendation 2003/361/EC	Referred to in the proposal or in the recommendations

The views expressed are the author's own.