Amnesty International UK Submission to the 2026 UK consultation “Growing up in the online world: a national conversation”

26 May 2026

Amnesty International UK (AIUK) is a national section of a global movement of over ten million people. Collectively, our vision is of a world in which every person enjoys all human rights enshrined in the Universal Declaration of Human Rights and other international human rights instruments. Our mission is to undertake research and action focused on preventing and ending grave abuses of these rights. We are independent of any government, political ideology, economic interest or religion.

This submission was produced in collaboration with Amnesty International’s global tech team on Children & Young People’s Digital Rights, who develop research, advocacy and campaigns to ensure young people can shape safe online environments and thrive in them.

Amnesty International UK

2nd Floor, Peter Benenson House,

1 Easton Street, London, WC1X 0DW

1 Benefits and risks of social media use for children and young people

Social media is a complex landscape for children's rights, offering meaningful opportunities for children to realize their rights, while also exposing them to serious risks. From a children's rights perspective, the harms are well-documented: Amnesty International's *Driven into Darkness* research shows that platform design itself can place children in harm's way.[1] TikTok's hyper-personalized "For You" feed can rapidly draw young users who show even minimal interest in mental health topics into "rabbit holes" of harmful content, including videos that romanticize or encourage depressive thinking, self-harm, and suicide. These findings highlight how platform design can exacerbate anxiety, depression, and self-harm risks for children and is fundamentally designed to maximize engagement by addicting users to the platform. At a population level, platform design choices can amplify misinformation, polarization, and even contribute to large-scale human rights abuses, as seen in international contexts such as Myanmar where online incitement on Facebook fuelled ethnic cleansing against the Rohingya Muslim population.[2]

On the other hand, social media also plays a vital role in enabling children and adults to exercise their rights. It is a key space for free expression and for accessing information, allowing young people to learn, communicate, and participate in public life. It supports freedom of peaceful assembly and association, enabling youth-led mobilization around shared human rights concerns in the UK and across the world, allowing young activists to connect, raise awareness and organize protests in support of diverse causes such as climate and racial justice, gender equality and anti-war protests. For many marginalized children, including LGBTI youth, online communities on social media are essential. Children and young people Amnesty works with repeatedly emphasized that, despite the risks, social media provides a crucial space for connection, learning, and protesting.

Given the complexities laid out here, a rights-respecting model of social media must both address the very real risks children face while not stripping away the spaces where they learn, connect, and exercise their rights. If we are to seek a safe, rights-respecting and vibrant future for social media, that means confronting the root cause of the problem: the business model that drives harmful design. Today's dominant platforms are engineered to maximize attention. The longer children and adults stay online, the more data these companies can harvest and the more advertising they can serve. This creates powerful incentives to deploy design features, including endless scrolling, hyper-personalized feeds and algorithmic amplification, that keep users addicted, even when those features expose children to harmful content or undermine their health and well-being. As long as this model remains intact, the harms will continue.

[1] Amnesty International, *Driven into the Darkness: How TikTok's 'For You' Feed Encourages Self-Harm and Suicidal Ideation* (Index: POL 40/7350/2023), 7 November 2023, https://www.amnesty.org/en/documents/POL40/7350/2023/en/

[2] Amnesty International, *The Social Atrocity: Meta and the Right to Remedy for the Rohingya* (Index: ASA 16/5933/2022), 29 September 2022, https://www.amnesty.org/en/documents/asa16/5933/2022/en; [2] Amnesty International, "UK: X's design and policy choices created fertile ground for inflammatory, racist narratives targeting Muslims and migrants following Southport attack", 6 August 2025, https://www.amnesty.org/en/latest/news/2025/08/xs-design-and-policies/

Amnesty International's research provides extensive evidence of the manifold risks and harms associated with this business model as well as proposing regulatory responses. The 2019 report *Surveillance Giants* called out Silicon Valley's surveillance-based business model as the root cause of many online human rights problems and fundamentally incompatible with human rights.[3] In our 2023 companion reports *Driven into Darkness* and *I Feel Exposed*, Amnesty demonstrated how TikTok's highly extractive business model, seamless hyper-personalisation and addictive design posed serious risks to the mental and physical health of children and young people.[4] This submission seeks to provide recommendations to the UK government on how to ensure children are safer online, focusing both on meaningful design changes – changes that make platforms less addictive, less manipulative, and less dangerous for children in the short term – as well as addressing the structural causes of Big Tech's assault on human rights.

To this end, Amnesty International makes the following **key recommendations**:

- To counter serious privacy and health risks associated with large social media platforms, regulators should ban the collection of intimate personal data, including data about their interests, emotional state or well-being which is inferred from a user's watch time and engagement for the purposes of 'personalizing' content recommendations and advertisements.
- Rather than using pervasive surveillance to adapt feeds to a user's interests, providers should be compelled to enable users to communicate their interests through deliberate prompts (for example, users could regularly be asked to enter or select specific interests if they would like to be served personalized recommendations) and all personalization must be based on users' freely given, specific and informed consent.
- Regulators must urgently ensure that businesses, including social media and AI companies as well as gaming providers, demonstrate the safety of their design choices before they are rolled out to children. Where companies have implemented design elements such as profiling-by-default, autoplay and infinite scroll on social media, sycophantic and anthropomorphic design in AI products or gambling-features in games without sufficient human rights due diligence in place, they should be required to withdraw these, until businesses are able to demonstrate their safety for minors, transparently, in line with international human rights standards and based on concrete evidence that is accessible to interrogation by independent researchers and auditors.

Please note that Amnesty International has put all allegations included in this submission to the companies named. To see the company responses please refer to the reports referenced in the footnotes where the responses are included throughout and in the annexes.

[3] Amnesty International, *Surveillance Giants: How the Business Model of Google and Facebook Threatens Human Rights* (Index: POL 30/1404/2019), 21 November 2019, https://www.amnesty.org/en/documents/pol30/1404/2019/en

[4] Amnesty International, *Driven into the Darkness: How TikTok's 'For You' Feed Encourages Self-Harm and Suicidal Ideation,* 7 November 2023 (previously cited) and Amnesty International, "*I feel exposed": Caught in TikTok's Surveillance Web* (Index: POL 40/7349/2023), 7 November 2023, https://www.amnesty.org/en/documents/POL40/7349/2023/en/

2 Children and young people's perspectives

In 2023, Amnesty International collected responses from 550 children and young people between the ages of 13 and 24 across 45 countries to better understand their lived experiences, concerns and attitudes towards social media.[5] Amidst praise for the diversity of ideas, users' creativity and opportunities for activism that young people find on social media, two major concerns stood out: the toll that harmful content and what many young participants described as "addictive" platform design take on young people's mental health and their feeling of powerlessness in the face of global companies' constant nudging to participate in a vicious cycle of personal data sharing and content consumption.

In more recent consultations with children and young people, we found a clear consensus: outright social media bans would detrimentally impact their human rights. Young people highlighted how crucial social media is for them - as a means of staying connected, especially for marginalized groups like LGBTI youth or young displaced people seeking community, as a vital source of information and learning, particularly around their rights and social issues, and as a space for their human rights activism.

> *"I remember when I first started to use the internet, the most meaningful part was being able to talk with people from around the world and have different perspectives on different topics. And it was really important to me not to feel alone."*
> *Paloma – teenage activist from Argentina*

> *"I spent more than a year off school before my GCSEs. I couldn't see any of my friends in person because my immune system was too weak. With social media, I was able to stay connected to what my friends were experiencing personally, and in the world outside..."*
> *18-year-old participant of a UK-based Youth Town Hall, hosted by FlippGen[6]*

> *"I publish videos on TikTok where I talk about human rights issues. And for me, a lot of the people that are following me are under the age of 16. And I wouldn't be able to reach them with this information if there was a social media ban."* Johanne – young activist from Norway

Underlining the systemic nature of social media harms and the limitations of policy solutions solely focused on younger teenagers rather than all users, 19-year-old UK-based digital rights campaigner Cosima shared:

> *"Young people are fed the message that you have to look perfect, be perfect, and make perfect choices, whilst what perfect is, shifts and shrinks by the day. Last year, harmful content in its persistence got to me - despite working to empower and safeguard other young people against these exact harms, I fell down the trap of eating disorder content online, and the pervasive harm that comes with it."*

[5] Amnesty International, ""We are totally exposed": Young people share concerns about social media's impact on privacy and mental health in global survey", 7 February 2023, https://www.amnesty.org/en/latest/news/2023/02/children-young-people-social-media-survey-2/
[6] Flippgen, Youth Town Hall: A youth-led report on creating a digital world for and by us, 2026, https://www.flippgen.com/the-report

These young people were acutely aware of the risks and harms posed by these platforms but they were also concerned that social media bans for teenagers had significant practical constraints, including with respect to what happens when teenagers over the age of the ban get access and are exposed to risks and the continuing risks for children who manage to circumnavigate the bans. Instead of outright bans, the children and young people we engaged advocated for more nuanced regulations that compel social media companies to change how their platforms operate. They stressed that they want states to stand up to Big Tech companies and prioritize children's safety by getting to the root cause of the issues and by enforcing design changes that pay respect to the autonomy and preferences of children and users more broadly, rather than simply excluding children from the platforms. Relatedly, they also underlined the importance of meaningfully including children and young people in this process as experts in relation to their own digital experiences and co-creators of positive alternatives to the current digital ecosystem.

3 Towards rights-respecting online platforms: restore online privacy and tackle addictive and deceptive design

3.1 Human rights risks and practical shortcomings of raising the age of teenage access

Amnesty International shares many of the concerns that motivated the launch of this consultation as we have extensively documented design risks and harms of social media platforms for both children and adult users.

The current focus in the public debate is on implementing age restrictions on teenagers' access to social media. However, this risks restricting children's rights whilst allowing Big Tech platforms to maintain a business model and design choices that run counter to children and adults' rights to a safe and privacy-respecting digital environment.

The only long-term, rights-respecting solution to ensuring children's rights and safety online is to tackle the harmful design of current Big Tech platforms, enforcing safety-by-design and user privacy through effective and properly enforced regulations. Robust safeguards are necessary to ensure social media platforms stop exposing users to harms through their relentless pursuit of user engagement and the exploitation of people's personal data.

While banning teenagers from social media may appear to offer a simple fix, it does not alter the platform design features that expose users of all ages to risk, nor does it resolve broader societal harms such as the spread of disinformation, polarization, or the role of social media in fueling human rights abuses.[7]

Bans affecting teenagers undermine children's rights: restricting access to social media limits young people's ability to seek and share information, participate in civic life, and associate with peers and communities - rights protected under the UN Convention on the Rights of the Child

[7] Amnesty International, "UK: X's design and policy choices created fertile ground for inflammatory, racist narratives targeting Muslims and migrants following Southport attack", 6 August 2025 (previously cited).

and defined in greater detail for the digital age in General Comment 25. Bans also risk disproportionately harming marginalized communities, such as LGBTI youth, who rely on online spaces for safety and support.

In practice, teenage bans are difficult to enforce as the Australian experience over the last six months clearly demonstrates: it is currently estimated that more than two thirds of children who previously had accounts on leading social media platforms still have access.[8] Teenagers are circumventing platforms' age restrictions or remain undetected as child users, highlighting clear issues with the effectiveness of age assurance tools. Conversely, critics warn that more effective age assurance tools may risk adverse impacts on the rights to privacy as well as potential chilling effects on the freedom of expression of both child and adult users (discussed in greater detail below).

Meanwhile, unenforced age restrictions risk leaving parents in the dark about their children's continued exposure to online risks, whilst the widespread failure of age restrictions is also already prompting governments to consider even more far-reaching measures like VPN restrictions, which raise serious privacy and human rights concerns (discussed in further detail below).

Finally, age restrictions on a select number of sites, as has been implemented in Australia, risk driving children toward unregulated platforms, games and AI applications that may pose similar serious risks linked to exploitative design and a lack of child rights due diligence.[9] On the other hand, proposals to age-gate a wide variety of websites risk disproportionate adverse impacts on the right to privacy, freedom of expression and children's right to education. Wider government powers to rate the risk of any given online platform or service and potentially censor access would pose serious risks to human rights and concentrate a lot of power in government hands, potentially even leading to authoritarian creep in the digital environment.

Alternative proposals to age restrictions include raising the age up to which parental consent is necessary to access social media platforms. Whilst such proposals recognize the importance of familial dialogue around children's online safety, they overlook the reality that parental control may not always be in children's best interest, e.g. in abusive family relationships, as well as potential discriminatory effects on children in the care of guardians other than their parents, practical concerns around some parents' ability to take up this task and privacy risks associated with the added data trail.

Ultimately, if nothing is fundamentally shifted in how these platforms are designed, then children over the age of the ban, children whose parents give consent (in case of a parental

[8] Guardian, *Two-thirds of under-16s with accounts on Instagram, Snapchat or TikTok kept access despite ban*, 31 March 2026, https://www.theguardian.com/australia-news/2026/mar/31/meta-tiktok-snapchat-google-under-investigation-australia-social-media-ban

[9] UNICEF, *Drawing a line in digital spaces: Age-based restriction of social media*, April 2025, https://www.unicef.org/media/170606/file/UNICEF%20policy%20note_age%20restrictions%20social%20media_FINAL%20%28003%29.pdf, Molly Rose Foundation, *Australia's social media ban – is it working?*, April 2025, https://mollyrosefoundation.org/wp-content/uploads/2026/04/MRF_Australia-Social-Media-Ban-Research_Briefing-April-26.pdf

consent-based age restriction), children who manage to circumvent age-gates, older teenagers and adults remain exposed to harmful design.

The government must ensure children's ability to enjoy the benefits of online interactions on platforms in ways that do not violate their right to privacy and put their health and well-being at risk.

The government won't be able to and shouldn't aim to lock children out of the internet in its entirety, yet the privacy and exploitative design harms associated with social media platforms are increasingly common across online services and games. We must thus address the systemic issues common to too many digital platforms and tools that haven't been designed with children's rights and safety in mind.

Amnesty International's caution is echoed by a broad spectrum of human rights experts, scholars and civil society organizations in the field of children's online safety. The Council of Europe's Commissioner for Human Rights stated in February 2026:

> *"Banning children's access to social media [...] shifts the responsibility for safety from the platforms that create the environment to the children who navigate it. States should require platforms to prevent and mitigate risks to children's rights by design and by default, and hold platforms accountable for failures.*
>
> *Given the pervasiveness of algorithmic systems, comprehensive regulation is essential. This includes ensuring algorithmic transparency and auditability, effective reporting and redress mechanisms, children's rights risk assessments, independent audits, and restrictions on targeted advertising. These obligations must be enforceable, subject to independent oversight, and supported by sanctions and liabilities that are effective deterrents [...] The source of harm is rooted in the design and incentives of the platforms. That should be the primary focus of regulation."*[10]

3.2 Addressing the systemic harms of hyper-personalized feeds and addictive design

Amnesty International's research has repeatedly highlighted the human rights risks and real-world harms of the engagement-based optimization of recommender systems based on the tracking and profiling of users through behavioural data, including where such data allows inferences about sensitive personal information. A human-rights respecting approach to social media regulation must therefore tackle the ubiquitous surveillance embedded in the current business model of Big Tech. Whilst these risks are not unique to child users, a permanent halt to behavioural tracking of children for the purposes of content recommendation on social media platforms is a critical step towards a safer online environment for children. Amnesty International has also been campaigning for a global ban on targeted advertising towards children

[10] Council of Europe Commissioner for Human Rights, *Regulate platforms, not children – Commissioner urges caution over social media bans*, 23 February 2026, https://www.coe.int/en/web/commissioner/-/regulate-platforms-not-children-council-of-europe-commissioner-for-human-rights-urges-caution-over-social-media-bans

since 2023, handing over a petition demanding such design changes from TikTok, signed by more than 170,000 global supporters, in 2025.[11]

Amnesty International's research on TikTok shows that the company has maximized the addictive qualities of design choices and engagement strategies employed by competing social media companies, incentivizing a race to the bottom between a small number of leading social media companies vying for the highest user numbers and engagement rates. TikTok has done this in spite of mounting scientific evidence of the serious risks associated with addictive use of social media especially for children and young people's health, including sleep and attention problems. The former US Surgeon General's advisory, published in 2023, noted that poor sleep does not just affect concentration and performance in school, it has also "been linked to altered neurological development in adolescent brains, depressive symptoms, and suicidal thoughts and behaviors".[12]

Our research with children and young people who were drawn into rabbit holes of depressive and self-harm-related content highlights the full extent of the risks associated with TikTok's addictive design and its hyper-personalized 'For You' Feed. Renewed Amnesty International research published in 2025 highlighted that the company continues to fail to mitigate the most harmful consequences of its engagement (i.e. profiling)-based recommender system, dragging children and young people who watch content related to mental health themes into feeds dominated by profoundly harmful content, years after this was first reported.[13]

To counter these serious privacy and health risks, large social media platforms must be compelled to cease collecting intimate personal data and drawing inferences from a user's watch time and engagement about their interests, emotional state or well-being for the purposes of 'personalizing' content recommendations and advertisements.

Amnesty International's recommendation is to ensure that access to and use of essential digital services and infrastructure are not made conditional on ubiquitous surveillance of children, young people or adult users. This will require enacting and enforcing comprehensive data protection laws in line with international human rights law and standards to prohibit targeted advertising on the basis of invasive tracking practices. Regulators should restrict the amount and scope of personal data that can be collected, strictly limit the purpose for which companies process that data and ensure inferences about individuals drawn from the collection and processing of personal data are protected from exploitation. The collection and use of inferred sensitive personal data (for example, recommendations based on watch time and likes which allow for inferences of sensitive information) to personalize ads and content recommendations must be banned.

Rather than using pervasive surveillance to adapt feeds to a user's interests, providers should be compelled to enable users to communicate their interests through deliberate prompts as the

[11] RTE, "Amnesty petition calls on TikTok to make platform 'a safer space'", 25 November 2025, https://www.rte.ie/news/business/2025/1125/1545650-tiktok-petition/
[12] US Surgeon General, Social Media and Youth Mental Health, May 2023, https://www.hhs.gov/surgeongeneral/priorities/youth-mental health/social-media/index.html
[13] Amnesty International, *Dragged into the rabbit hole: New evidence of TikTok's risks to children's mental health* (Index: POL 40/0360/2025), 20 October 2025, https://www.amnesty.org/en/documents/POL40/0360/2025/en/

default option (for example, users could regularly be asked to enter or select specific interests if they would like to be served personalized recommendations) and all personalization must be based on users' freely given, specific and informed consent.

Infinite scroll, autoplay, push notifications:

Adding to the addictive effect of hyper-personalized content recommendations, social media platforms now employ a set of exploitative design choices aimed at maximizing time spent on their platforms and putting their users into auto-pilot mode. Amnesty International's research on TikTok, a driver of many of these changes across the industry, highlighted infinite scroll, autoplay, push notifications and "likes" as key elements of this strategy. All of these design choices aim to minimize friction, condition users to lose track of time and get lost in a never-ending stream of content.

Additional concerns have been raised by young people recently consulted by Amnesty International in relation to "streak" functionalities on social media platforms which reward users with a 'streak' for regularly (usually daily) logging on to the platform or performing a certain task such as messaging a contact.

Neurologist Servane Mouton further explained the mechanisms and effects in a 2025 interview with Amnesty International:

"Behaviours on social networks are similar to addiction, with this difficulty in getting off, the encroachment on other activities, the obsession with thoughts in a loop around having access to them. Missing [them] when you're not there. The economic model of social networks is based on stimulating the short-term reward system so that we go there as often as possible, for as long as possible. It is a bit like chasing after dopamine shots that will immediately produce the sensation of pleasure and make us want to go back."

Importantly, Amnesty International's research highlighted that older teens and young adults, also reported struggling to combat the addictive lure of social media, and that the available user tools, at least on TikTok, were ineffective at curbing the problem.

Despite these known risks, internal documents revealed in the context of recent US lawsuits show that TikTok's leadership explicitly warned product teams against implementing changes that would reduce "stay time" of "minors and excessive users" by more than 5%.[14]

With regards to Meta, the Knight Georgetown Institute's analysis of internal company documents stated that "internal understanding of risk has not necessarily translated into product change" citing one plaintiff's expert reviewing internal documentation who suggested that "Meta's awareness of 'Problematic Use' [...] is 'High,' while their 'Product investment' is 'Mid-Low.'"[15] This is even though Meta's internal research found that more than a tenth of global Facebook

[14] Knight Georgetown Institute, *Measuring Risk: What EU Risk Assessments and US Litigation Reveal About Meta and TikTok*, 12 February 2026, https://kgi.georgetown.edu/research-and-commentary/measuring-risk-what-eu-risk-assessments-and-us-litigation-reveal-about-meta-and-tiktok/

[15] Knight Georgetown Institute, *Measuring Risk: What EU Risk Assessments and US Litigation Reveal About Meta and TikTok* (previously cited), p. 21.

users self-report problematic use and that the company is aware that Meta's youngest users are at greatest risk.[16]

"Likes":

Adolescent research participants also commented on the perceived role of the 'like' function, common to all leading social media platforms, in exacerbating insecurities and keeping their eyes fixed on the screen. They explained that they felt compelled to keep checking their account to monitor how their posts fare in comparison with others' posts. Numerous psychological studies and – as the "Facebook Papers" revealed – Big Tech's own in-house research have documented the toxic effect of social media driven social comparison on the mental health of adolescents.

Psychologists interviewed as part of Amnesty International's research and prior independent research also pointed to negative impacts of design features encouraging social comparison between children and young people and their peers, such as the "heart" or "like" feature, on the mental health of adolescents. For instance, Maëlle and Édouard, 18 and 17 at the time of Amnesty International's research in 2025, reflected on the risks of the "like" or "heart" button in relation to videos discussing self-harm and suicidal ideation:

"When I liked the video, it was to encourage, to give strength, to let the person know that they can be well...But I later realized that it might actually mean to the person, "I'm being encouraged to self-harm." (Maëlle)

"The online support can also be interpreted as 'go ahead and kill yourself'." (Édouard)

The abovementioned experiences of young research participants highlight the most catastrophic effects of such design choices. Whilst the level of harm may not be equal across all children and young people, the evidence nonetheless paints a sufficiently clear picture to warrant urgent action under the precautionary principle. Leading social media companies have failed to fulfil their responsibility to respect human rights and have failed to implement effective child and human rights due diligence in line with international business and human rights standards as well as binding regulations such as the EU's Digital Services Act. Even where the latter introduced specific reporting requirements on companies' risk assessment and mitigation, leading businesses have thus far failed to provide transparent and concrete evidence that would substantiate their claims of adequate risk mitigation.[17]

[16] Trial Report of Mitch Prinstein at paragraph 53, In Re: Social Media Adolescent Addiction/Personal Injury Products Liability Litigation, No. 4:22-md-03047-YGR (Northern District of California December 11, 2025), cited in Knight Georgetown Institute, *Measuring Risk: What EU Risk Assessments and US Litigation Reveal About Meta and TikTok* (previously cited).

[17] DSA Civil Society Coordination Group, "Initial analysis of the first round of risk assessment reports under the EU Digital Services Act", March 2025, https://cdt.org/wp-content/uploads/2025/03/RA-Report-Assessment-Report.pdf

Regulators must therefore urgently ensure that businesses demonstrate the safety of their design choices before they are rolled out to children. Where companies have implemented design elements such as profiling-by-default, autoplay and infinite scroll without sufficient human rights due diligence in place, they should be required to withdraw these, until businesses are able to demonstrate their safety for minors, transparently, in line with international human rights standards and based on concrete evidence that is accessible to interrogation by independent researchers and auditors. If companies are unable to counter the increasing evidence of certain design choices aimed at maximizing engagement being inherently unsafe, regulators should consider prohibiting their deployment towards children or – where appropriate – all users.

3.3 Deceptive and exploitative design

Deceptive design choices undermining informed consent, nudging children and adult users into excessive use or facilitating economic exploitation are found across children's online experiences, often in the form of so-called "dark patterns", across social media, websites, online games and AI chatbots and companions. Regulators should consider the increasingly interconnected nature of these previously distinct online experiences as social media platforms integrate AI chatbots and companions and games replicate social functionalities, integrate AI personalization and copy profiling and ad targeting techniques.

Online tools and platforms, including social media platforms, use "dark patterns", often also referred to as "deceptive design", to encourage users, including the many millions of children and young people around the world that use these platforms, to sign up to their services, agree to their terms of service and, in the process, give access to huge amounts of personal data. The huge troves of data and the insights drawn from them are ultimately exploited for profit.

Legal scholars and subject experts have repeatedly criticized the "click-wrap"' nature of these "contracts"' that users, including children and young people, commonly accept without reading. The impossibility of reading, let alone absorbing, the terms and conditions of such contracts has been highlighted in the past by academics, who in 2008 found that "a reasonable reading of all the privacy policies one encounters in a year would require 76 full workdays".[18] Given the increased use of and reliance on online services and platforms as part of daily life, it is reasonable to assume that this figure would now be higher. The use of these "dark patterns" and the difficulty of reading and understanding the terms of service means that any consent granted by children and young people cannot be considered to be genuinely free and informed consent.[19] The opaqueness of these terms of service impact on the right to privacy, as they directly impact a user's ability to have control over their own information.

Further undermining the user's autonomy and conscious decision-making, social media platforms have also been observed to implement time management tools that give clear visual preference to the option that allows the user to stay on the app rather than logging off. Dark patterns are

[18] The calculation was made by two professors at Carnegie Mellon University in the USA. See Shoshana Zuboff, The Age of Surveillance Capitalism, 2019, pp. 48-50.
[19] Amnesty International, "*I feel exposed": Caught in TikTok's Surveillance Web*, 7 November 2023 (previously cited).

prohibited under the EU's Digital Services Act and partially addressed under the UK's Children's Code and yet they remain common across social media platforms and other digital environments accessed by children, pointing to clear failures to enforce existing laws and regulatory instruments.

To disrupt manipulative design practices, regulators should require that companies provide clear information to their users about the purpose of collecting their personal data from the start and that they do not further process it in a way that is incompatible with this purpose or their responsibility to respect human rights. Social media companies should be compelled to provide age-appropriate explanations to children of their terms of service. These should use clear and simple language, provide transparent information throughout the user process and not only at the beginning, and provide clear explanations throughout children's online experiences of user control choices, settings and features.

Academic research into design risks in the gaming sector highlights the urgent need to broaden out regulatory awareness of exploitative design risks across sectors. The prevalence of "free-to-play" games underpinned by the same profiling/surveillance-based business model as social media platforms has been shown to have led to a surge in exploitative design choices targeting child and adult gamers with the aim of extending time spent, encouraging in-game purchases and maximizing ad revenue.[20] Particularly high-risk and exploitative practices include the profiling and targeting of ads at children and gambling-like features such as loot boxes. Age-appropriate design may require a mixture of design feature prohibitions with respect to children as well as changes in default settings and requirements in relation to user autonomy and parental controls.

3.4 Emerging risks in the context of AI chatbots and companions

Independent reporting on AI chatbots, which are now increasingly being embedded into online platforms, raises concerns that these systems' designs and functionalities could pose risks to children's rights. These include concerns that AI chatbots may encourage unhealthy forms of attachment, dangerous behaviours and excessive use with potential negative effects on users' mental health and that they may lack sufficient safeguards to address the risks inherent in interacting with minors experiencing mental health issues.[21] Other concerns raised relate to the privacy of user interactions, the creation of deepfakes and child sexual abuse material (CSAM) and the inappropriate use of sexual language in chats with minors.[22]

---

[20] Working Group on Gaming and Regulation, Feedback on the European Commission's Digital Fairness Act, October 2025, https://bhr.stern.nyu.edu/publication/feedback-on-the-european-commissions-digital-fairness-act/

[21] Common Sense Media, "Social AI Companions", 16 July 2025, https://www.commonsensemedia.org/articles/social-ai-companions-0, Cathy Mengying Fang and others, "How AI Chatbots Affect Our Social and Emotional Wellbeing: New Research Findings", 21 March 2025, https://www.media.mit.edu/publications/how-ai-and-human-behaviors-shape-psychosocial-effects-of-chatbot-use-a-longitudinal-controlled-study/, Guardian, *Teenager died after asking ChatGPT for 'most successful' way to take his life, inquest told,* 31 March 2026, https://www.theguardian.com/society/2026/mar/31/teenager-asked-chatgpt-most-successful-ways-take-life-inquest-told

[22] Reuters, *Italy fines OpenAI over ChatGPT privacy rules breach,* 20 December 2024, https://www.reuters.com/technology/italy-fines-openai-15-million-euros-over-privacy-rules-breach-2024-12-20/, Internet Watch Foundation, *AI CSAM Report 2026*, 3 April 2026, iwf-ai-csam-report-2026.pdf, Reuters, *Meta's AI rules have let bots hold 'sensual' chats with kids, offer false medical info*, 14 August 2025, https://www.reuters.com/investigates/special-report/meta-ai-chatbot-guidelines/

Through their integration into toys, social media and messaging platforms, AI chatbots are rapidly expanding their reach into children's lives. Without adequate data protection safeguards, they risk amassing huge amounts of knowledge about a child's life and most personal thoughts from a toddler age, turning them into profit as training data. AI companions marketed as virtual "friends", "partners" and even mental health aids entice children and young people to share their most intimate thoughts and experiences. The surveillance-based business model traditionally associated with social media platforms is thus crossing a new frontier, with potentially far-reaching consequences for the right to privacy and freedom of thought. Anthropomorphic design choices aimed at maximizing engagement manipulate children into assigning human traits to machines, opening them up to manipulation and exploitation. Studies with children and adolescents show that younger users can 'overtrust' AI and may be more prone than adults to assign human traits to AI agents based on anthropomorphic design choices.[23]

Several widely reported suicide cases are pointing to severe mental health risks associated with the use of AI companions or chatbots for children and young people with mental health issues.[24]

Current UK legislation fails to address most of these risks and even where instruments such as the Children's Code offer at least partial responses, widely available GenAI products fail to comply. Children must be recognized as a protected group necessitating additional safeguards where GenAI products are targeted at or accessible to them, including a ban on the manipulation and exploitation of children and the creation of AI-generated or AI-adapted CSAM. A risk-based approach should categorize products accessible to children as high-risk, necessitating mandatory human rights due diligence including child rights impact assessments, offering public and transparent evidence of businesses' risk assessment and mitigation.

Protected groups, including children, as well as civil society stakeholders must be meaningfully involved in the risk assessment and mitigation process. As discussed above, lessons learnt from the risk assessment and mitigation framework under the EU's Digital Services Act should prompt regulators to require that companies assess the safety of design features before launch and urgently intervene to reverse the roll-out of high-risk design choices and features unless or until companies provide transparent and substantiated evidence of their safety that is open to probing by independent researchers and auditors. This includes sycophantic and anthropomorphic design as well as chatbot prompts that nudge users to reactivate conversations or that amount to emotional manipulation.

[23] Jaemarie Solyst and others, *Children's Overtrust and Shifting Perspectives of Generative AI*, June 2024, https://arxiv.org/pdf/2404.14511, Judith Stanja and others, "Children's and Adolescents' Anthropomorphic Conceptions of Social Robots and Chatbots – A Systematic Literature Review", Koli Calling '25: Proceedings of the 25th Koli Calling International Conference on Computing Education Research, Article No.: 9, Pages 1-10, https://dl.acm.org/doi/full/10.1145/3769994.3770002

[24] New York Times, Can a Chatbot Named Daenerys Targaryen Be Blamed for a Teen's Suicide?, 23 October 2024, https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html, CNN, Parents of 16-year-old sue OpenAI, claiming ChatGPT advised on his suicide, 27 August 2025, https://edition.cnn.com/2025/08/26/tech/openai-chatgpt-teen-suicide-lawsuit, Guardian, *Teenager died after asking ChatGPT for 'most successful' way to take his life, inquest told,* 31 March 2026 (previously cited).

4 Age assurance

All age assurance tools involve the collection and processing of personal data to some extent. The nature of age assurance means that both children and adults must provide their personal data to either prove they are above a certain age in the case of a child, or over the age of 18 in the case of an adult. This may present challenges to the right to privacy for all users, which can then also have a knock-on effect on other rights. Given these risks, we recommend that greater emphasis be put on design choices that platform providers can make to increase safety-by-design as opposed to implementing age assurance/verification and increasing data collection. Where age assurance and age verification tools are implemented, it is important that they are designed in privacy-preserving ways.

As has been emphasized by Joseph Cannataci, former UN Special Rapporteur on the right to privacy, without the conditions of a private life and private spaces, the full potential of the individual and their fundamental right to human dignity are compromised, and therefore privacy is critical to the development of personality.[25] With regard to age verification, he noted that where verifiable identity documents are required, this raises concerns around security and prescriptive approaches that place the emphasis on age-gating parts of platforms over design choices that can be made to make the platforms safer for children.[26]

The development of age assurance tools raises further concerns regarding a user's more expansive data trail in relation to their online activity and its potential misuse by corporate and state security agencies. The erosion of online privacy carries further risks in relation to potential chilling effects on freedom of expression, in particular for journalists, activists, and those wishing to express political dissent amidst broader concerns surrounding surveillance and policing powers in relation to peaceful protests in the UK.[27] The risk of an added data trail applies to adults as well as children. It raises the concern of profiling by the private sector or state actors who may be able to track the places children and adults go online through the trail left by age verification tools.

This may lead to companies either over-restricting content and access to online spaces for children for compliance reasons, or conversely the use of children's data for commercial purposes such as advertising. The widespread use of age verification online may also exacerbate structural discrimination within society, shutting out already marginalized groups unable to meet the requirements of age-assurance methods whether due to a lack of documents, economic disadvantages that preclude users from having their own devices or by excluding them from spaces where they fail AI-powered age estimation tests trained on data that lacks sufficient diversity. Cybersecurity experts have also repeatedly warned that identity verification and age

[25] OHCHR, Statement by Mr. Joseph A. Cannataci, Special Rapporteur on the right to privacy, at the 31st session of the Human Rights Council, 9 March 2016, https://www.ohchr.org/en/statements-and-speeches/2017/02/statement-mr-joseph-cannataci-special-rapporteur-right-privacy-31st

[26] OHCHR, Statement by Mr. Joseph A. Cannataci, Special Rapporteur on the right to privacy, 30 November 2020.

[27] Amnesty International, "UK accused of complicity as world faces most dangerous moment for human rights - annual report", 21 April 2026, https://www.amnesty.org.uk/latest/uk-accused-of-complicity-as-world-faces-most-dangerous-moment-for-human-rights/

assurance tools fail to provide adequate safeguards against data theft.[28] It is important to consider and protect child and adult users alike from these potentially adverse impacts on human rights, in light of the claimed positive impact of age assurance tools. While the intent behind age assurance tools is to safeguard children online by mitigating their exposure to potentially harmful content and services and protecting them from online child sexual exploitation and abuse, there are concerns that it will be possible to circumvent any age assurance tools implemented and that they will provide a false sense of security, where users may wrongly assume that the verified age claimed by an individual is accurate. Thus, there are concerns around the ability of such tools to achieve this aim, which should be taken into account when considering the necessity – and therefore lawfulness - of any such tools that pose risks to rights and may not in fact prevent children's exposure to harmful content or features. Amnesty International therefore determines that the focus of regulatory interventions should instead be on ensuring safety-by-design.

5 VPNs as critical privacy tools

Amnesty International strongly opposes measures to restrict, age-gate or effectively prohibit the use of Virtual Private Networks (VPNs), whether directly or by requiring VPN providers to perform age assurance on their users or to block access to specified content. Such measures would fail – in the context of social media bans for children - the tests of necessity and proportionality required under international human rights law, would cause collateral harm to the much larger population of legitimate users, and would not achieve the child protection objectives the consultation invokes.

VPNs are a fundamental component of digital security infrastructure used by businesses, public institutions, journalists, human rights defenders, remote workers and ordinary users. The UK's National Cyber Security Centre (NCSC) recommends VPN use as a core element of device security.[29]

Public Wi-Fi networks - ubiquitous in cafés and transport hubs - present documented, serious security risks and 43% of UK users have experienced security incidents on public networks.[30]

VPNs are security-critical infrastructure to be protected not weakened: policies that disrupt them create new risks exploitable not only by authoritarian states but by criminal actors. Research presented by the Citizen Lab identified novel exploits that could be used against major VPN implementations.[31]

---

[28] Panda Security, Over one billion customer records belonging to IDMerit users left unprotected online, 18 March 2026, https://www.pandasecurity.com/en/mediacenter/customer-records-idmerit-unprotected/, Cybernews, EU age verification app can be hacked in 2 minutes, claims security expert, 21 April 2026, https://cybernews.com/security/eu-age-verification-app-hack/

[29] National Cyber Security Centre (NCSC), 'Virtual Private Networks', Device Security Guidance v2.1, reviewed May 2025, https://www.ncsc.gov.uk/collection/device-security-guidance/infrastructure/virtual-private-networks

[30] Zimperium, "Travel Is Up and So Are the Risks: 5 Million Public Unsecured Wi-Fi Networks Exposed", 2025, https://zimperium.com/blog/travel-is-up-and-so-are-the-risks-5-million-public-unsecured-wi-fi-networks-exposed

[31] Benjamin Mixon-Baca and Jeffrey Knockel (Citizen Lab), 'Attacking Connection Tracking Frameworks as used by Virtual Private Networks', Privacy Enhancing Technologies Symposium (PETS), Bristol, 16 July 2024, https://citizenlab.ca/vulnerabilities-in-vpns-paper-presented-at-the-privacy-enhancing-technologies-symposium-2024/

Beyond individual security, VPNs are a critical tool for accessing information under conditions of censorship and surveillance. In Iran for example, recent research shows that most users relied on VPNs as their primary means of reaching major platforms including following renewed government restrictions.[32] For journalists, activists and human rights defenders operating in repressive environments, VPNs are frequently the difference between safety and exposure.

Any restriction on VPN use engages the rights most commonly exercised online: freedom of expression, access to information, freedom of association, and the right to privacy. Under international human rights law, including Human Rights Council Resolution 38/7 (2018) on the promotion, protection and enjoyment of human rights on the Internet - any restriction on the exercise of these rights must be lawful, pursue a legitimate aim, and be both necessary and proportionate to that aim. We explain below why VPN restrictions to enforce social media bans fail both the necessity and proportionality tests.

5.1 Necessity: the link between VPNs and social media circumvention is weak

The necessity test requires that a restriction must be necessary and the least intrusive means capable of achieving the legitimate objective. On this measure, the case for VPN restriction in the context of children's social media access is not sufficient because VPN restrictions are an intrusive measure (see section 5) and yet, the evidence that they would be effective in this context is lacking.

VPN usage in the UK did increase substantially following "highly effective age assurance" requirements coming into force in July 2025, settling on around 800,000 users at the end of the year,[33] but there is no evidence of substantial numbers of children using VPNs to circumvent age checks.[34]

Despite the increases in VPN usage, the Age Verification Providers Association has reported an additional five million age checks daily since the new duties came into force, suggesting the great majority of users are engaging with age checks rather than circumventing them through VPN use.[35]

In Australia, a survey of teenagers by the Molly Rose Foundation in April 2026 found that 61% of 12–15-year-olds who previously held accounts on restricted platforms still had access despite the social media ban being in place since late 2025. Critically, 70% of those children said that platforms simply failed to spot their existing accounts. Only approximately one in twenty children was using a VPN.[36]

---

[32] Iran International, 'Targeting platforms', February 2026, https://www.iranintl.com/en/202602196792

[33] HM Government, 'Growing up in the online world: a national conversation' (consultation document), Box 3.3

[34] "New data shows no rise in children's VPN use after the introduction of online age checks" Katie Freeman-Tayler | 4th December, 2025 https://www.internetmatters.org/hub/research/data-shows-no-rise-childrens-vpn-use-amid-online-age-checks/

[35] Mishcon de Reya / Louise Schofield, "Online Safety Act: VPNs and age verification - what the House of Lords debate reveals", citing Age Verification Providers Association data, https://www.mishcon.com/news/online-safety-act-vpns-and-age-verification-what-the-house-of-lords-debate-reveals

[36] Molly Rose Foundation, 'Australia Social Media Ban Research Briefing', April 2026 (previously cited).

Even in the most extreme context, child sexual abuse material (CSAM), where there is unambiguous legal consensus on the need for action, the data does not support VPN restriction as an effective response. The Childlight Searchlight 2025 report found that of over 853,000 CSAM page requests blocked over two days globally only 2.5% were from users using VPNs, and only 15% used any anonymisation service whatsoever.[37]

Notably, the Age Verification Providers Association's has also stressed that "there is no need to ban VPNs to make age assurance effective." Technical and behavioural indicators can reveal enough information even with a VPN.[38] A VPN masks an IP address; it does not mask a device history, a social graph, or years of geotagged content.

Given the evidence that VPNs are not being widely used for circumventing age gates where they are in effect, and that banning them may not be necessary for the intended purpose, age gating VPNs is unlikely to meet the necessity test under international human rights law.

5.2 Proportionality: technical ineffectiveness and serious collateral harm

Even if residual necessity could be demonstrated, VPN restrictions would likely fail the proportionality test on a number of reinforcing grounds.

Technical ineffectiveness

The Knight Georgetown Institute's Age Assurance Technical Assessment states plainly that "all age assurance systems are vulnerable to circumvention. It is not technically feasible to... prevent all minors from accessing restricted content... without also blocking large numbers of adult users." Restricting VPNs is explicitly identified as having 'widespread negative security and privacy consequences... without preventing circumvention by determined users who would migrate to the Tor anonymity network, decentralised privacy tools, or other censorship-resistant technologies.[39] Tor (The Onion Router)[40] is a widely-used freely available decentralised anonymity network specifically engineered to resist censorship.

Collateral harm to legitimate users

Anonymous access to social media is only useful in limited cases, as the personal profiles can negate the advantages of anonymity, but can be very important for human rights defenders and other targeted groups or individuals. Blocking or age-gating VPNs would weaken everyday digital privacy for everyone and increase risks for these users who rely on VPNs for safety. A joint statement signed by 438 security and privacy scientists from 32 countries in March 2026 warned that VPN restrictions would harm a very large population of legitimate users of these

[37] J. Stevenson and others, "Access Denied: How Blocklists are Thwarting Attempts to View CSAM", in Childlight, Searchlight 2025 - Who Benefits? Shining a Light on the Business of Child Sexual Exploitation and Abuse, Edinburgh: Childlight, 2025.
[38] Age Verification Providers Association (AVPA), "AVPA Responds to Criticisms from Computer Scientists", https://avpassociation.com/thought-leadership/avpa-responds-to-criticisms-from-computer-scientists/
[39] Rescorla, E., Arnao, Z., & Cooper, A. (2026, January). *Age assurance online: A technical assessment of current systems and their limitations*. Knight-Georgetown Institute. https://kgi.georgetown.edu/wp-content/uploads/2026/01/Age_Assurance_Online_Technical-Assessment_Report_KGI.pdf
[40] https://www.torproject.org/

technologies - remote workers, journalists, security researchers, human rights defenders - in order to achieve stated child protection benefits.[41]

We believe that banning VPNs would be disproportionate and once set down that route any attempts to reduce circumvention such as by blocking Tor and similar technologies, would also be disproportionate.

5.3 Fully blocking VPNs risks setting a dangerous precedent

Some commercial VPN providers' exit nodes are already blocked by certain online services on security grounds, their IPs having been used for some unlawful activity, but this is a blunt process that produces significant false positives and requires continuous maintenance. Scaling this to a comprehensive government-mandated UK-wide block would require infrastructure investment and a degree of traffic surveillance that is disproportionate to the benefits.

The UK has an established legal and technical infrastructure for blocking websites at ISP level, comprising DNS blocking, IP address blocking, URL filtering, and hybrid systems such as BT's Cleanfeed mechanism, which combines IP blocking with deep packet inspection (DPI) for URL-level filtering.[42] This infrastructure underpins both High Court injunctions (used for copyright infringement and illegal streaming) and Ofcom's enforcement powers under the Online Safety Act but it does not stop VPN use.

There is a further technical point not acknowledged in the consultation document. VPN services are primarily user-friendly packaging of open-source technologies such as OpenVPN, WireGuard and IPsec, which are in turn built on fundamental networking and cryptographic protocols forming part of the basic architecture of the internet itself. These protocols cannot meaningfully be banned without disrupting banking transactions, health records and government communications. Blocking VPN traffic at the network level, would require the UK to deploy a very concerning censorship infrastructure.[43]

Even short of an outright ban, measures to require VPN providers to perform age checks on their users, or to block access to specified content for users who cannot verify their age, would introduce a structural mechanism with serious long-term risks. The UN Special Rapporteur on the Freedom of Opinion and Expression, in his foundational 2015 report on Encryption and Anonymity, concluded that encryption and anonymity tools 'deserve strong protection' as they 'enable individuals to exercise their rights to freedom of opinion and expression in the digital age', and that 'blanket prohibitions fail to be necessary and proportionate.' He recommended that states 'avoid all measures that weaken the security that individuals may enjoy online' and 'refrain from making the identification of users a condition for access to digital communications and

[41] 'Joint Statement of Security and Privacy Scientists and Researchers on Age Assurance', open letter, 438 signatories from 32 countries, signatures closed 9 March 2026, https://csa-scientist-open-letter.org/ageverif-Feb2026
[42] On UK ISP blocking architecture see Open Rights Group project Blocked https://www.blocked.org.uk/about
[43] Tamlin Magee, 'Could the UK ban VPNs if it wanted to?', Raconteur, 6 August 2025, https://www.raconteur.net/technology/could-the-uk-ban-vpns-if-it-wanted-to

online services.'[44] A UK framework requiring VPN providers to log user identities and selectively block content could set a dangerous global precedent.

Imposing conditions on mainstream VPN providers could also have the unintended consequence of driving users, including children, to less regulated and more data-exploitative services. Free VPNs frequently monetise through data collection, and one popular free extension with over 100,000 downloads was discovered to be silently screenshotting users' pages.[45]

5.3 VPN circumvention of social media and website blocking

This consultation is framed primarily around social media, but the discussion of VPN restrictions in the UK has not been limited to that context. The Children's Commissioner described VPN circumvention as a 'loophole that needs closing',[46] and House of Lords debates have raised VPN accessibility in relation both to social media access and to the availability of specific harmful websites.[47]

These include VPN usage to access a suicide forum that has been warned to restrict UK visitors and may eventually be blocked.[48] Banning VPNs to prevent circumvention of lawfully mandated website blocking in the UK may appear more justified, particularly in cases such as preventing suicide, but it is still deeply problematic and likely ineffective.

Amnesty International's concerns about this broader question go beyond the immediate consultation question on VPNs, but they are engaged directly by any proposal to extend network-level blocking.

6 Use of secondary legislation and ministerial powers

We believe that the powers inserted into the Online Safety Act 2023 by Section 70 of the Children's Wellbeing and Schools Act 2026 are too broad, raising serious concerns that are separate from and additional to the substantive arguments about any intervention.

Although the current government may want to use these powers responsibly following the results if this consultation, any such commitments are not binding on a future government empowered by the Act.

This provision grants the Secretary of State a sweeping power to "amend or repeal primary legislation" - forcing providers of "specified internet services" to prevent or restrict access by

[44] UN Special Rapporteur on Freedom of Opinion and Expression (David Kaye), 'Report on Encryption, Anonymity, and the Human Rights Framework', A/HRC/29/32, 22 May 2015. Available at: https://digitallibrary.un.org/record/798709
[45] Leonard Bernardone, 'VPNs won't save teens from social media ban', Information Age / Australian Computer Society, 11 December 2025, https://ia.acs.org.au/article/2025/vpns-won-t-save-teens-from-social-media-ban.html
[46] Children's Commissioner Dame Rachel de Souza, statement reported in: 'Children's Commissioner urges action to stop children using VPNs', Computing, 2025, https://www.computing.co.uk/news/2025/legislation-regulation/children-s-commissioner-urges-action-to-stop-children-using-vpns
[47] Hansard, House of Lords, 'Children: Age Verification and Virtual Private Networks', 4 December 2025, https://hansard.parliament.uk/lords/2025-12-04/debates/4ECC433F-7DB4-43ED-845B-687DFE5D8B50/ChildrenAgeVerificationAndVirtualPrivateNetworks
[48] Molly Rose Foundation, 'Missed Chances, Lost Lives', October 2025, https://mollyrosefoundation.org/wp-content/uploads/2025/10/MissedChancesLostLives.pdf

children.[49] There are also similar powers to change the age of consent for children in data protection laws.

The substantive decisions of this consultation about which services are in scope, which age thresholds apply, and what compliance steps are required will be made by the government through so-called "Henry VIII powers" that bypass the full parliamentary scrutiny that primary legislation commands.

After a short debate, Parliament can only vote to accept or reject the regulations but not to amend them.[50] The Hansard Society, in a briefing published during the passage of the Bill, expressly identified the section 70 online safety provisions as raising "serious constitutional concerns about parliamentary scrutiny and accountability."[51]

No limitations in scope

The substantive scope of these ministerial powers is exceptionally broad. The term "specified internet services" can encompass virtually any online service offered to users in the United Kingdom. There is no restriction in the primary legislation to social media, or even to user-to-user services. The "Relevant child" is set by the regulation, potentially anyone under 18. Although the purpose is protecting children from harm, the legislation doesn't restrict these powers to services presenting identified risks to children.

By contrast, Australia's Online Safety Amendment (Social Media Minimum Age) Act 2024 - the main reference in the consultation process - applies to a narrower defined category of "age-restricted social media platforms"; and expressly exempts messaging services, online gaming, professional networking, educational platforms, and health services.[52] The UK's proposed architecture contains no equivalent statutory exemptions. These would be left entirely to the ministerial discretion of whoever holds the relevant office in the future.

7 Overlap and insufficient enforcement of the existing regulatory framework

The legislation does not adequately address the extent to which the objectives it identifies are already met - or could be met without new primary legislation - by the existing regulatory framework, and where it may overlap or clash with those.

Section 81 of the Data Use and Age Assurance Act amends Article 25 of UK GDPR to require that services "likely to be accessed by children" implement data protection by design and by

[49] Section 70(7)(f), Children's Wellbeing and Schools Act 2026 (legislation.gov.uk/ukpga/2026/21/section/70/enacted): regulations under new section 214A of the Online Safety Act 2023 "may amend or repeal primary legislation."
[50] Supplemental memorandum to the Delegated Powers and Regulatory Reform Committee https://bills.parliament.uk/publications/65342/documents/8002
[51] Hansard Society, 'Last-minute powers and limited scrutiny: Parliament and the risks of consigning online safety law to delegated legislation' (2026), hansardsociety.org.uk/publications/briefings/online-safety-delegated-legislation-social-media-ai
[52] Online Safety Amendment (Social Media Minimum Age) Act 2024 (Cth): definition of 'age-restricted social media platform' and express exemptions for messaging services, online gaming, professional networking, educational platforms, and health services.

default, with explicit recognition of children's heightened vulnerability.[53] The Online Safety Act 2023 separately requires user-to-user services and search services to conduct children's risk assessments, implement age-appropriate safety measures, and - for the most harmful categories of content - apply highly effective age assurance to exclude under-eighteens.[54]

The ICO's Age Appropriate Design Code imposes binding standards on information society services "likely to be accessed by children", operating under the data protection regime.[55] These include requirements for privacy by default; data minimisation and others. The Children's Code applies across a similarly broad range of internet services as the potential scope of Section 70, and the ICO has already fined organisations like Reddit for failing to implement strict age checks.

Taken together, existing legal instruments under data protection and online safety already constitute a substantial and insufficiently enforced framework of child protection obligations for online services, apart from a social media ban. The consultation should address with greater rigour why this existing framework is insufficient on its own terms, rather than proceeding with the assumption that broad new executive powers are required.

8 Changing the age of digital consent

The legislation creates a power for government ministers to raise the digital age of consent - the age at which children may themselves provide consent under UK GDPR for the processing of their personal data by information society services - from its current level of 13 to a higher threshold, potentially 16.[56] As the consultation document explains, "most social media services that set a minimum age do so at 13. This aligns with the age of digital consent."[57]

We are concerned that changing the age is a technical corollary of the proposed access restriction, not an independently evidenced child welfare policy, and could have wider impacts on the rights of children.

If social media services are required to exclude under-sixteens, and if the access of thirteen-year-olds currently involves their digital consent, then the age of consent must be raised to be coherent with the ban. The consultation itself acknowledges that raising the age could widen digital inequalities and limit access to beneficial services including educational technology.[58] These are serious risks to which the consultation provides no substantive response.

[53] Section 81, Data Use and Age Assurance Act, amending Article 25, UK GDPR.
[54] Online Safety Act 2023, sections 11–27 (children's risk assessments and safety duties); section 68 (highly effective age assurance for content harmful to children, including pornographic content and content promoting self-harm, suicide, or eating disorders).
[55] ICO, Age Appropriate Design Code (2021), issued under section 123, Data Protection Act 2018. The Code applies to information society services "likely to be accessed by children" (defined as persons under 18). See ICO, About the Children's Code (ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code).
[56] Article 8, UK GDPR; section 9, Data Protection Act 2018. The current threshold is 13 years in the United Kingdom.
[57] DSIT, "Growing up in the online world: a national conversation" (CP 1528, March 2026), p. 17.
[58] DSIT consultation, pp. 21–22.

Changing the age of digital consent for adolescents between thirteen and sixteen needs to consider how the complexities of that specific developmental period shape the rights of those children in line with the concept of children's evolving capacities. The concept of evolving capacities affirms that as children acquire greater competencies, the need for protection diminishes, and their capacity to engage more fully in decisions that affect them increases.[59] This concept has been extensively examined in relation to legal capacity for consenting to medical treatment, where the foundational *Gillick* principle is that parental rights to unilaterally decide progressively diminish - and ultimately terminate - as the child acquires sufficient understanding and intelligence to make the decision.[60] At sixteen, legal capacity is presumed to be present; below sixteen, it must be assessed contextually but cannot be assumed to be absent.

Section 9 of the Data Protection Act 2018 already reflects this principle: it exempts the processing of children's data for preventive or counselling services from the parental consent requirement, erring on the side of caution to avoid preventing adolescents from seeking help confidentially. The new Section 8ZA(5) inserted by the 2026 Act would preserve this carve-out in the event of age changes.[61]

This exemption does not resolve the broader tension with the principle of evolving capacities under international human rights law and the graduated, capacity-based model of adolescent autonomy that English law has recognised since *Gillick*. Even if the change was aimed at services relying on consent as their lawful basis, it would have a broader impact. For example, a 15-year-old with full understanding of a service's data practices would nonetheless require parental authorisation to agree to specific data processing. The proposals risk creating new barriers to the health, sexual health, counselling services, and LGBTQ+ community information that young people are most likely to need to access independently of their parents.

9 Amending Section 393 of the Communications Act 2003: transparency as a prerequisite for human rights accountability in online regulation

We believe that any reform of the online regulatory regime must address how Section 393 of the Communications Act 2003 (CA 2003) prevents public scrutiny of Ofcom's exercise of its powers under the Online Safety Act 2023 (OSA).

Section 393 CA 2003 prohibits the disclosure of information obtained by Ofcom "with respect to

[59] UNICEF, "Growing with Rights Understanding and supporting the evolving capacities of the child", 2026, https://www.unicef.org/innocenti/media/12651/file/UNICEF-Innocenti-Growing-Rights-report-2026.pdf
[60] Gillick v West Norfolk and Wisbech Area Health Authority [1986] AC 112, per Lord Scarman at 188–189: "the parental right to determine whether or not their minor child below the age of sixteen will have medical treatment terminates if and when the child achieves a sufficient understanding and intelligence to enable him or her to understand fully what is proposed."
[61] New Article 8ZA(5), inserted into UK GDPR by the Children's Wellbeing and Schools Act 2026: "In paragraph 1, the reference to information society services does not include preventive or counselling services."

a particular business" in the exercise of its statutory powers.[62] The Online Safety Act 2023 extended Ofcom's remit without revisiting Section 393, with the result that a confidentiality provision designed for the market regulator of the telecommunications sector was carried wholesale into online safety regulation, a completely different context. The OSA vests Ofcom with extraordinary powers with direct and significant consequences for the rights to freedom of expression, privacy, and access to information.

The EU's Digital Services Act provides a relevant contrast with the UK OSA: it imposes affirmative transparency obligations on very large online platforms, requires regulatory decisions to be made publicly available in non-confidential form, and has established a dedicated whistleblower mechanism enabling protected disclosure of inside information about platform practices.[63]

Section 393 is being applied to restrict online safety accountability through FOIA

The ability of the public and civil society to scrutinise how Ofcom's powers to regulate the internet are exercised is central for the democratic accountability of such interferences with fundamental rights.

The Freedom of Information Act 2000 was designed to provide a balance of such scrutiny and the commercial confidentiality of information that companies share with public bodies such as Ofcom. The private interests of those businesses are balanced with public interest considerations for disclosure. FOIA also has stronger absolute exemptions to protect trade secrets - such as algorithms – that override any public interest considerations.

Section 393 has been lethal to Ofcom's transparency and accountability over who is influencing them and on what evidence - because Section 44 FOIA provides an absolute exemption for information whose disclosure is prohibited by other legislation like S 393 - bypassing any public interest considerations.[64] Clean Up the Internet has established that over 50% of FOI requests relating to the OSA are being refused or partially refused on this basis.[65]

Ofcom has refused to disclose any details of meetings between its senior leadership and major platforms including Google, X, Meta, and TikTok; in some cases, declining even to *confirm or deny* whether any such meetings had taken place. Other refusals include whether companies fined by Ofcom under the OSA have paid their fines.

---

[62] Section 393, Communications Act 2003. The provision prohibits disclosure of information "with respect to a particular business" obtained by Ofcom "in exercise of a power conferred by" the legislation under which Ofcom operates, including - following the commencement of the Online Safety Act 2023 - its online safety functions.

[63] European Commission, Digital Services Act: transparency obligations on very large online platforms and search engines, including public reporting requirements, access for vetted researchers, and a dedicated whistleblower tool for the submission of inside information about platform practices (digital-strategy.ec.europa.eu/en/policies/dsa-whistleblower-tool).

[64] Section 44, Freedom of Information Act 2000 provides an absolute exemption for information whose disclosure is prohibited by another enactment. Unlike most FOIA exemptions, it does not require a public interest balancing exercise.

[65] Clean Up The Internet FOI response (September 2025): at the time of that request, over 50% of FOI requests relating to the Online Safety Act had been refused or partially refused on the basis of Section 393 CA 2003 (via the Section 44 FOIA absolute exemption). https://www.cleanuptheinternet.org.uk/post/ofcom-legal-challenge

In a particularly worrying case, s393 was used to reject a FOIA request of the "small number" of stakeholders who Ofcom said had persuaded them to insert a significant change into the OSA Illegal Content Codes of Practice (*Measure ICU-C2)*.[66] This secrecy means not being able to examine and question the arguments that weakened the obligations requiring platforms to swiftly take down illegal content, which - after lobbying - Ofcom made conditional on technical feasibility.[67]

Commercial confidentiality provisions shield from scrutiny how a powerful public regulator is influenced by global technology corporations in making decisions affecting millions of people's rights. These are some of the largest corporations in the world, with substantial resources to participate in regulatory processes through private channels. Meanwhile, civil society organisations representing communities most affected by those decisions are denied the most basic information about how those processes are conducted.

In September 2025, the Information Commissioner's Office upheld Ofcom's position, accepting that Section 393 CA 2003 prohibited disclosure of the regulator's engagement with tech companies.[68] Campaign groups are progressing a judicial review of Ofcom's approach,[69] but Section 393 creates a structural secrecy regime at the heart of the OSA's implementation that requires a legislative remedy.

[66] Ofcom stated this change followed feedback from "a small number" of stakeholders, using evidence from WhatsApp as an example, but later refused to identify any other stakeholders, or disclose what evidence they had provided. The rest of Ofcom's response included many examples from the submissions of tech companies to the consultation. (https://www.ofcom.org.uk/siteassets/resources/documents/online-safety/information-for-industry/illegal-harms/volume-2-service-design-and-user-choice.pdf?v=390978 para 2.40)

[67] Ofcom, Illegal Content Codes of Practice, Measure ICU-C2: the final published version inserted the qualification "unless it is currently not technically feasible for them to achieve this outcome" into the obligation to swiftly take down illegal content. (https://www.ofcom.org.uk/siteassets/resources/documents/online-safety/information-for-industry/illegal-harms/illegal-content-codes-of-practice-for-user-to-user-services-24-feb.pdf?v=391889)

[68] ICO Decision Notice IC-403651-J1L1 (September 2025), upholding Ofcom's refusal to disclose information about its regulatory engagement with major platforms under the Online Safety Act 2023. https://ico.org.uk/media2/eycdrh1n/ic-403651-j1l1.pdf

[69] Clean Up The Internet v Information Commissioner, First-tier Tribunal (General Regulatory Chamber), listed for hearing 26 February 2026. Clean Up The Internet, "Ofcom faces legal challenge", cleanuptheinternet.org.uk/post/ofcom-legal-challenge; the Movement for an Open Web has applied to join as a party. See Preiskel and others, "Regulatory pressure mounts as MOW applies to join CUTI in its fight over Ofcom's transparency", preiskel.com/regulatory-pressure-mounts-as-mow-applies-to-join-cuti-in-its-fight-over-ofcoms-transparency