**Growing up in the Online World: consultation response**

---

**Summary**

We welcome the Government's consultation on children's online safety and recognise the speed at which it has been produced and the work undertaken by the Department for Science, Innovation and Technology (DSIT) while it has been opened to ensure wide engagement with parents and carers, children and young people, and civil society and academic experts. We have particularly valued the willingness of DSIT officials to engage directly with the Network's central team and with our partners to understand their perspectives and ideas before the consultation closes.

The proposals contained below have been shared with DSIT Ministers and officials in various written briefings and meetings during the course of their development in recent months; blog posts relating to the bulk of the material below have been published online during the course of the consultation.

This submission brings all of this material together in one place for ease of reference and constitutes our response to the consultation. We have not submitted a separate response to the individual questions via the consultation portal.

**About the Online Safety Network**

The OSA Network informs, coordinates and supports effective, ongoing civil society engagement and advocacy with policymakers, regulators and Parliamentarians on matters relating to online safety and AI harms. We develop and deliver effective solutions that work across different harm domains. Interests represented in the Network include: child protection, terrorism, extremism, violence against women and girls, suicide and self-harm prevention, mental health, hate speech and online abuse, fraud and scams, animal cruelty, mis- and disinformation, harms to democracy and threats to our information environment. The Network is led by Maeve Walsh (Director) and Prof Lorna Woods OBE (expert legal adviser) and continues work undertaken at Carnegie UK during the passage of the Online Safety Bill.

**Our approach to the consultation**

This consultation response from the Online Safety Act Network focuses on four areas:

- The policy context, online safety objectives and potential outcomes

- The proposed interventions relating to children's access to online services and potential age-gating of features and functionalities
- The proposed interventions relating to AI chatbots
- Wider interventions to strengthen the Online Safety Act and deliver sustainable change

We recognise that the consultation format sets out a series of proposed interventions relating to children's online safety and asks for respondents' views, often via tick-box preferences, and we fully understand the reason the Government has taken this approach, and also the ease with which this allows feedback from the wider public to be sought and collated.

From a policymaking perspective, this significantly narrows the scope of the ideas and proposals under consideration and also narrows the scope for external experts, whether from civil society, academia, industry or other professions, to critique the options available or offer alternatives. Given the pressures the Government is under (political, parental, time, resources), this is an entirely understandable approach. It will not necessarily lead to the most effective or sustainable outcomes, however - for children and young people, parents or indeed for the Government.

We have chosen therefore to propose alternatives which, we believe, will deliver that more effective and sustainable impact on children's online safety and further strengthen the foundations of the Online Safety Act. We set these out below.

**Note:** we have not had the capacity in the time available to develop a detailed written response to the proposals relating to compliance and enforcement (highly effective age assurance) or preparing children for a digital future (media literacy, parental support). We would refer DSIT to the analysis and recommendations in the written submissions from expert organisations in our Network, such as 5 Rights Foundation and Internet Matters respectively, which we support.

**Our recommendations**

Our key recommendations are:

- **Strengthen the Online Safety Act as per our 10 Point Plan**. Without doing this, any new interventions to improve children's online safety will only be half-measures and the challenges arising from Ofcom's implementation and enforcement of the OSA remain. Please see the attached PDF: Strengthening the Online Safety Act: our 10 Point Plan.
- **Adopt our Safety by Design Code of Practice for immediate implementation by Ofcom**. The OSA does not need to be amended for this to happen, but strengthening it to make its adoption a requirement and bring in a clear definition of "safety by design" is also recommended. Please see the attached PDF: Safety by Design Code of Practice.
- **Bring in comprehensive, risk-based duties on AI chatbots** to keep children - and all users - protected from design-based harms. Please see our attached PDF with an amendment to bring this into force, along with our research paper on the harmful design of these products.

The following documents are attached to this consultation

- Written evidence from Prof Lorna Woods to the Science, Innovation and Technology Committee on social media age restrictions
- Our new Safety by Design code
- Our 10 point Plan to Strengthen the Online Safety Act
- The text of a proposed standalone duty for AI chatbots
- A research brief on AI chatbots

**Policy context, online safety objectives and potential outcomes**

We published a blog covering some of these themes on 7 April.

Context

There is no doubt that something needs to be done to improve children's online safety and hold tech companies, particularly social media platforms, more strongly to account. The successful grassroots parental campaigns are a symptom of that, particularly driven by parents of younger children who want to delay their offspring's exposure to social media (or, in some cases, digital devices in general) and increasingly by health professionals, who are dealing with significant numbers of children and teenagers presenting with problems that have social media access and/or exposure as a contributing factor.

The arguments for and against a "ban" have been well-rehearsed[1], and our expert legal adviser, Prof Lorna Woods, gave oral and written evidence to the Science, Innovation and Technology Committee's inquiry earlier this year, focusing particularly on scope and effectiveness.[2] Many thousands of words have been written and spoken about the proposals - from proponents and opponents - since January.

But it is important, we believe, to consider what policy question a "ban" is trying to solve. This is a necessary first step to allow an assessment of whether a ban (or access restriction) is the right approach and how it would then need to be designed, implemented and enforced to ensure it delivers that outcome. This question may not matter for campaigners calling for "something to be done" - and indeed, the pressure that this seemingly simple solution ("ban children") has created has already delivered an important result in opening up this opportunity to consider the next steps for protecting children online. But it does matter if the Government's consultation is asking people to respond "yes" or "no" to it.

Framing

We think the consultation should have been clearer on whether the proposals (bans or other measures) are intended to improve children's online and/or offline wellbeing (eg improving childhood experiences and outcomes)[3] or protect them from harms primarily associated with online environments (eg exposure to inappropriate content or activity).

[1] The views of many of our Network partners can be found here: https://www.onlinesafetyact.net/analysis/under-16-social-media-bans-views-from-our-network/

[2] Written evidence to the SIT Committee: https://www.onlinesafetyact.net/documents/1487/Evidence_to_SIT_Committee_inquiry_into_social_media_age_restrictions.pdf; transcript of oral evidence: https://committees.parliament.uk/oralevidence/17322/pdf/

[3] In the opening information on p5, it states "the government is seeking views on its proposals to ensure children have **enriching** digital lives"; the Secretary of State in her foreword talks about what makes a "great childhood"; and the introduction talks about "time spent away from screens in healthy real-world environments".

The introduction switches between both these perspectives and ends with a "simple ambition: to make sure that our children's lives online are as safe and healthy as they are offline, while equipping them with the confidence, skills and support to thrive in a world transformed by rapid technological change".

In the next chapter, it states that we "want to deepen our understanding of how children engage with digital technologies in their daily lives to inform future government work", with a number of paragraphs describing various uses of technology by children which "enrich" their lives, including education and learning, creativity, gaming, communication and language learning. This is then followed by a canter through the problems linked to "patterns of use" (screentime, sleep deprivation, displacement of offline activities) which may be affecting "critical parts of health [sic], happy childhoods where every child achieves and thrives" (p15) and the "features and functionalities" within platforms which "can exacerbate harm for children".

The Government's approach is disappointingly limited here in respect of one of the biggest challenges around digital technologies - whether enriching or otherwise: the business models. In the introduction, the Government refers to the fact that the consultation is "considering options to address concerns that the business model of certain apps .. designed to keep users online for longer" (p10) and there is a reference in chapter 1 to "concerns about business models that are built around driving engagement". But there are no corresponding questions that respondents can engage with on problems from "business models", except obliquely via tick box responses on the harm relating to lists of "addictive" or "persuasive" features and functionalities designed to keep children online for longer.

Despite the high-level overview of some of the positives and negatives from digital technology, we strongly believe that the simple ambition set out in the introduction cannot be delivered upon by this consultation. It assumes that children's "offline" lives are a quantifiable benchmark for "safe" and "healthy" and can be measured as such in isolation from their "online" lives. A "great childhood" comprises both online and offline dimensions, but the consultation is silent on the interventions that the Government would like to make to improve the "offline" aspects of childhood safety and health in tandem with the online impacts. It also assumes that digital technologies - products, services, devices - are either "positive" or "negative" depending on the content they serve up or the interactions they offer children, when the reality is very nuanced.

Excessive screentime at the expense of other activities can have a detrimental impact on a child's development and physical health[4], regardless of whether the screen use is related to "enriching" or educational content, just as exposure to harmful activity or content can have a detrimental impact on a child's safety or their emotional or psychological wellbeing, regardless of whether it is a one-off occasion or the result of cumulative exposure. One-size-fits-all interventions - in particular bans or age

[4] Muppala et al: "Effects of Excessive Screen Time on Child Development: An Updated Review and Strategies for Management" (2023): https://pmc.ncbi.nlm.nih.gov/articles/PMC10353947/; Khumukcham & Singh (2023): "The hazards of excessive screen time: Impacts on physical health, mental health, and overall well-being"; https://pmc.ncbi.nlm.nih.gov/articles/PMC10852174/

restrictions aimed at a limited number of platforms and services - will not address the complexity of what a "good childhood" means, or improve "wellbeing", particularly if the Government is unable to describe what those ambitions look like.

In short, a consultation document, rooted in evidence and structured coherently, that delivered on this simple ambition would have taken longer than the six weeks the Government had to produce it. We welcome the fact that the document opens up a number of long-overdue conversations about children's social media use, the design of social media platforms and the levers that the Government might have to address these issues. But we have remained concerned throughout the consultation that it risks doing no more than just testing the temperature on what works and what doesn't within the bounds of that narrow scope.

The subsequent Parliamentary pressures and the Government's legislative commitments in response have further narrowed the options open to them: the amendments to the Children's Wellbeing and Schools Act commit the Government to action on access restrictions (a ban) or age-gating features and functionalities, with no legislative levers now open for any other routes[5]. Crucially, given the political context in which it was announced and launched, the Government does parents and children a disservice implying that this document can do more than this. Nearly two and a half years after the Online Safety Act received Royal Assent, it is therefore ultimately disappointing that the first consultation on how to improve it is so limited.

Regulatory and policy context

The limitations of the consultation are even more disappointing because the Government's faith in the OSA as the "foundation" and "strong baseline" (p10) for the way forward - and its promise that "any new measures will align with this existing framework" - fails to acknowledge the problems within the structure of the Act itself. At her first Parliamentary Questions following her appointment as Secretary of State for Science, Innovation and Technology, Liz Kendall answered a question on the effectiveness of the Online Safety Act: "I will be paying close attention to what is working and will not hesitate to go further if necessary," she said; and, referring to the addition of self-harm material to the list of priority offences in the OSA, "I hope this shows the House my determination to take all necessary steps on this issue."[6]

That was on 10th September 2025: six months on from that, this consultation gives respondents no opportunity to put suggestions on what further necessary steps she might take to strengthen the

[5] We have published analysis of the Government's proposed legislative powers (the Henry VIII amendments) here https://www.onlinesafetyact.net/analysis/osa-amendments-and-henry-viii-clauses/ This narrowing of routes for meaningful action on online safety has been further compounded by the lack of any vehicles to bring forward online safety legislation in the King's Speech. See our recent blog here: https://www.onlinesafetyact.net/analysis/what-the-king-didnt-say-the-online-safety-shaped-hole-in-the-governments-legislative-agenda/

[6] https://hansard.parliament.uk/Commons/2025-09-10/debates/D0CDD497-3A2B-4916-8BA7-6DF0DD6EF818/OralAnswersToQuestions

"foundations" of the Act - for all users - if it is to deliver on its promise. We have worked extensively in the past 12 months to develop a package of recommendations - our 10 Point Plan for Strengthening the Online Safety Act - to do just that and have submitted this previously to the DSIT Secretary of State, the Minister for Online Safety and DSIT officials. The recommendations are backed by 24 organisations in our Network and would deliver cross-harm improvements, if implemented.[7] We have attached a PDF of the Plan to this submission for completeness. If the Government is not going to bring forward a further consultation on online safety any time soon, then these are the bare minimum required to ensure the existing regime works as intended.

The Government in its consultation notes that "all regulatory regimes need to remain agile" and that acting "swiftly" is necessary in the "fast-moving world of technology". There is no reference in the consultation as to where the Government - or Ofcom, its regulator - might have been more agile and swift in the two and a half years since the OSA received Royal Assent. We set out later in this submission how many of the measures the Government are consulting on here could have already been implemented by Ofcom and what this delay tells us about the pressures the Government are now under to act "swiftly".

There are welcome references to "safety by design", but again, starting from an assumption that this has already been delivered under the existing framework eg with references to the OSA "already" calling for services "to take a safety by design approach to protect children" from harmful content (p15), and an factually incorrect statement in chapter 3 that "Any new approaches to enforcement will build on and complement existing approaches to keeping children safe online, *including existing safety-by-design duties* and the Protection of Children Codes of Practice". [8] As we have consistently argued, Ofcom's interpretation of this requirement has been lacking[9]. If this consultation shifts the regulator towards a greater understanding of what "safety by design" means, that will be welcome; to aid this, we have published a code of practice on safety by design to demonstrate what good looks like and we discuss this further below.

[7] https://www.onlinesafetyact.net/analysis/strengthening-the-osa-our-10-point-plan-for-government/
[8] While there is a statement in section 1 of the Act that "Duties imposed on providers by this Act seek to secure (among other things) that services regulated by this Act are safe by design", this is not the same as a "safety by design" duty.
[9] https://www.onlinesafetyact.net/analysis/safety-by-design/

**Proposed interventions relating to children's access to online services and potential age-gating of features and functionalities**

We published a blog covering some of these themes on 7 April.

We are concerned that, as a result of political pressures, the Government has boxed itself into a corner - compounding the narrow terms of the consultation with legislative commitments that require its outcomes to be defined and progressed within two months of the consultation closing[10]. While the urgency for action is not in dispute, the way in which the Government has now pre-judged the consultation and set up a binary either/or mechanism for its implementation is short-sighted. Even if a political decision has been made that these two options are the right and only response, there are many other additional measures - many of them which could be implemented more quickly - which should also be on the table. We set out one such below but first set out some analysis of the Government's preferred options..

The Government's preferred options

*"Bans" or access restrictions*

The first set of options, now backed up by the amendment to the Children's Wellbeing and Schools Act, relate to where to set the age of access to social media services: respondents who support a legal requirement for a minimum age of access are asked whether that should be "at least 16" (as a yes or no answer) or "lower than 16", with options given as 13, 14, 15 or "other". A further question asks for views on the impacts of setting the minimum age "higher than 13". Note that these options refer only to social media services, despite the consultation in other sections talking about other services and products, including gaming.

Ofcom has recently published its report on tech firms' responses to its call for details on how they were implementing a number of measures under the OSA.[11] In its press release, the regulator says: "Despite acknowledging the importance of minimum age policies, none of the companies with a minimum age of 13 on their services convinced us that they are currently enforcing them effectively and the impact is clear. Our latest research shows that 84% of children aged 8–12 are still using one of the top five reaching online services (YouTube, Facebook, TikTok, Instagram and Snapchat) despite a minimum age of

[10] https://bills.parliament.uk/publications/66126/documents/8234 The Government's amendment - requires that the Secretary of State, following its "Growing Up in an Online World" consultation **must "**by regulations make provision requiring providers of specified internet services—

(a) to prevent access by children of or under a specified age to specified internet services which they provide, or to specified features or functionalities of such services;

(b) to restrict access by children of or under a specified age to specified internet services which they provide, or to specified features or functionalities of such services."

[11] https://www.ofcom.org.uk/siteassets/resources/documents/online-safety/protecting-children/project-mercury/update_tech-firms-responses-to-our-call-for-action-to-protect-children.pdf?v=418243

13. Current online safety laws do not explicitly require services to keep underage children off their platforms by using robust age checks, although the Information Commissioner's Office can take action in this area under data protection law. We have today written to the Secretary of State to advise that, should Government and Parliament wish Ofcom to be able to force firms to enforce minimum age policies effectively, this would need a clearer basis in online safety legislation."

This gives the Government a clear evidential basis to act on enshrining a minimum age of access in law - whether the existing standard minimum of 13 across most social media platforms, or whether a higher age of 16 ("delaying" access, or "banning"). It is a relatively rare occurrence for the regulator to advise the Government so publicly on recommended changes to legislation, too. But Ofcom's position that the law does "not explicitly require services to keep underage children off their platforms by using robust age checks" is not entirely true. The Online Safety Act contains duties that require category 1 services to enforce their terms of service; given that minimum age of access criteria are included in ToS, it would seem reasonable to assume that the enforcement of those minimum age criteria is within Ofcom's power. (A caveat here, however, is that - as we set out in our 10 Point Plan - there is nothing to stop regulated services "rolling back" protections set out in their ToS and we recommend that the OSA is updated to remove that loophole.) The problem for Ofcom is that - due to the delays in publishing its register of categorised services - it has not yet consulted on these duties, or a raft of other duties related to category 1 services. That consultation comes in July, meaning that Ofcom's enforcement of the duties is unlikely to start until early 2028[12]. It is therefore something of an irony that Ofcom has judged that a law that is already in place to allow it to enforce against companies that are breaching it will take longer to enforce than asking the Government for a new law to be enacted.

We would ask the Government to be clear with the public - and particularly parental campaigners - when it responds to Ofcom's request and/or when it publishes its report to Parliament on taking forward the legislation on age restrictions (whether set at 13 or higher) exactly how long they expect this new law to be enforced not just the deadlines to which they are working to lay the regulations to bring it into force. Our assessment is that, even with a minimum 12-month deadline for regulations to be laid (July 2027) and assuming that Ofcom has enough clarity about the detail of the regulations to start its consultation on the guidance to implement those regulations as soon as they are laid, it is likely to be a further 16 months until those restrictions are enforceable: December 2028. Even if we assume that the political imperative to act reduces those timeframes and/or the Government is clear that when the regulations come into force, companies need to take steps to comply immediately, it is likely that a 12-month grace period for compliance will need to be introduced (as per the Age Appropriate Design Code), taking us to July 2028.

[12] We are calculating this based on the length of time (16 months) it has taken previous OSA duties to be consulted upon by Ofcom, finalised and brought into force. For example, the illegal content duties: consultation (November 2023), statement (December 2024), in force (March 2025); the child protection duties: consultation (March 2024), statement (May 2025), in force (July 2025); additional safety measures: consultation (June 2025); statement (tbc autumn 2026).

*Restrictions based on features and functionalities*

The second set of options that are linked to the commitment to regulate via the Children's Wellbeing and Schools Act are on restrictions to services based on features and functionalities and the consultation lists a number of examples: livestreaming, ability to send and receive images and videos containing nudity, location sharing, stranger-pairing and disappearing messages. The consultation questions ask whether these functionalities (or others, for respondents to specify) should be age-restricted and, if so, what the preferred minimum age should be. A further question asks to what extent respondents agree or disagree with the statement "restricting children's access to these features/functionalities would provide for a safer online experience for children", and views are also sought on what the impacts would be if age restrictions on these features were brought in.

We would like the Government to be explicitly clear in its response that action on these features and functionalities could already have been taken by Ofcom under the OSA. With the exception of the ability to send and receive images and videos containing nudity, **all of these features were flagged by Ofcom as potentially harmful to children** in both their draft register of risks in May 2024 and in the final register of risks, published in April 2025. Yet the regulator chose not to introduce any measures relating to these features in either their illegal harms or children's codes of practice - which respectively came into force last March and last July. They have since consulted on adding livestreaming as an additional safety measure with the details on that due to be incorporated into a further iteration of the codes later this year. (Please see our table setting out the gap between Ofcom's identification of other risky features and functionalities and corresponding mitigation measures to address the risk across both their sets of codes for more analysis of these gaps.[13])

While it is welcome that the Government is in effect taking matters into its own hands here, and proposing to bring in regulations that specify the need for age-gating for individual features, the missed opportunity over the past two and a half years since the OSA came into force is regrettable. It is also, as we set out below, likely to be compounded by what is now a rushed, piecemeal approach to playing catch up - picking out a limited number of individual features for post-hoc intervention (eg children can't access them) rather than taking a more systemic approach to ensure safety for children across the entirety of the service and its design. How often does the Government expect that it will need to update this list and bring forward new regulations to list new features or functionalities that have not as yet been identified as causing harm, or have not even been developed and designed into services?

Moreover, age-gating services that include these 'risky functionalities' as part of their model, will do nothing to stop those design choices being made in the first place - and adult users, older teenagers (16 and up) as well as under-16s bypassing age verification, will continue to experience harm at the hands of Big Tech.

[13] https://www.onlinesafetyact.net/documents/298/OFCOM_CODES_-_UPDATED_MEASURES_TABLE.pdf

A further set of questions, also in scope of the updated regulations, asks for views on "persuasive design" features associated with "addiction", with infinite scroll, autoplay, affirmation functions and alerts and push notifications singled out for consideration, along with personalised algorithms.

Respondents are asked which ones of those are "particularly 'persuasive'", which features should be age-restricted and at what minimum age. Additional questions ask for views on setting daily screen time limits for individual apps or restricting overnight access (so-called "curfews"). In its recent Parliamentary statements, the Government has also made clear that considerations on these latter measures are in addition to the either/or ban or age-restrict features and functionalities decisions.

We wish to acknowledge that this welcome focus on features and functionalities demonstrates an understanding from the Government that platforms are making design choices that are actively harmful to their users to generate more engagement, and, subsequently, more profit. This has been extensively researched and evidenced by civil society and survivors alike, from algorithmically recommended self-harm and suicide content[14], targeted advertising and filter bubbles[15], through to sexually explicit search suggestions for children[16] and the psychological harm caused by deceptive design choices[17].

But again, there have been many missed opportunities under the existing OSA framework to address these issues. In the consultation document, the Government observes that "the DSIT Select Committee noted in their recent report, 'Social media, misinformation and harmful algorithms', thinking about these questions through the lens of the business model can be helpful". What the Government doesn't mention is that last summer it rejected every single one of that Committee's recommendations to amend the OSA or take action on some of the structural and systemic problems the Committee had identified that cause harm - and on which the Government is now rushing to act. We note that the Chair of the Select Committee, Dame Chi Onwurah, made these points in the Committee's response to the consultation:

> "The government accepted almost all the conclusions in our July 2025 Report on 'Social media, misinformation and harmful algorithms', but almost none of our recommendations. We urge government to revisit them now and that should include bringing forth new legislation. For example, we concluded that social media companies are not merely platforms but curators of content, the amplification and spread of which can have serious impacts. They should be regulated as such. Meta told us that their algorithm can manipulate engagement with content, reducing engagement by downranking content by up to 80-90%. The algorithms that these platforms create must not be allowed to blindly promote any content that drives engagement while relying on individuals to report each item of harmful or illegal content. The volume of such content is growing and is not being blocked effectively, or at all, in many cases."[18]

[14] https://mollyrosefoundation.org/wp-content/uploads/2025/08/proof3_PervasivebyDesign.pdf
[15] https://riskybydesign.5rightsfoundation.com/recommendation-systems
[16] https://counterhate.com/research/x-rated/
[17] https://riskybydesign.5rightsfoundation.com/recommendation-systems
[18] https://committees.parliament.uk/publications/53021/documents/296367/default/

Mindful of the comparisons with the Australian social media ban - which the Government were keen to emphasise when they launched the consultation but which is only applicable to a very small number of services - respondents are also asked which type of services they think the restrictions should apply to. A separate section looks at options relating to AI chatbots, which we discuss further below.

A coherent, future-proofed, comprehensive way forward

Regardless of the Government's commitment to take one or other of the options above, we strongly urge that this is not the limit of its proposals. There are three particularly urgent changes included in our 10 Point Plan for Strengthening the Online Safety Act to recommend here. Most urgently - and requiring only minimal, targeted amendments to the legislation - three of our proposals are directly relevant to the failure by Ofcom to have acted earlier on harmful features and functionalities and are interlinked: introducing a requirement for regulated services to address all the risks identified on their services; removing the "clear and detailed" and "technically feasible" criteria for code measures; and removing with the "safe harbour" provision, which limits companies' compliance with the Act to the measures in the codes. These issues need to be addressed urgently to prevent further gaps emerging in OSA implementation and enforcement in the coming years. We can provide the Government with text to deliver these amendments, if helpful.

Secondly, the Government must ensure that in Ofcom's forthcoming consultation on the phase 3 duties for categorised services under the Online Safety Act, which include enforcement of terms of service, regulated services are required to enforce the minimum age of access that already exists on most services. The Government says in its consultation that there is "No current minimum age for accessing social media set in law" and Ofcom's recent report has played that back to them, asking for the law to be changed. As we note above, what it doesn't mention is that the standard industry minimum age for access on most services - set out in their terms of service and providing the basis for questions on a user's age when they set up an account - is 13. The general industry failure to enforce their services' standard minimum age of 13 - which the European Commission has recently taken action against Meta on under the DSA - has been known about for years. If Ofcom fails to include the enforcement of platforms' existing minimum age as a requirement in the forthcoming duties on categorised services then there is little hope for serious action from them if the Government decides instead to introduce measures to either restrict under-16s from accessing platforms entirely or introducing age-gating for individual features.

Finally, the introduction of a safety by design code - which is also one of our 10 Point Plan proposals, is directly relevant to the approach the Government is trying to take by identifying individual design features for post-hoc interventions. It was also recommended by the SIT Committee in their recent letter and is a key finding from the extensive research on the implementation of children's online safety

regulation published recently by Steve Wood[19]. The Ada Lovelace Institute, in their major recent report on children's digital lives, also recommended the adoption of our code of practice.[20]

With reference to some of the challenges with Ofcom's narrow approach to implementation above, a safety by design approach moves the focus of the regulation onto outcomes. One of the aspects that has been particularly frustrating for civil society with regard to the selection of measures in the existing OSA codes of practice is Ofcom's reliance on "evidence" that those measures work before they can recommend them, rather than putting the onus on the regulated services to either a) develop their own measures that are proportionate and appropriate to the risks identified on their own platforms; or b) take upstream design steps to design out those risks in the first place. Evidence of the kind required by Ofcom is only available if that measure already exists: which become a self-limiting brake on continuous improvement and innovation and, indeed, leaves the way open for services to remove existing measures that they may already be implementing if they are not included in the codes. By taking a feature-by-feature approach here, the Government risks falling into a similar trap: attempting to define for services what they must do (while those services sit on their hands and wait for the direction of travel), rather than holding them to account for the existing design of the features that already doing that cause harm. Unless the Government already has an extensive evidence base on the particular features that they wish to age-gate, then we would argue that an underpinning safety-by-design approach is all the more important to deliver improvements in the meantime.

A safety by design code: what it is and how it works

Safety by design[21] has emerged as a central principle in digital regulation, reflecting a shift toward tech accountability that requires digital services to assess and mitigate risks to users from the earliest stages of product development and throughout the entire lifecycle of the product or service. The Online Safety Act (OSA) explicitly references the need for user-to-user services to be 'safe by design' on the face of the legislation in section 1(3)[22]. Furthermore, the Secretary of State for the Department for Science, Innovation and Technology (DSIT) sets out safety by design as a key priority in their Statement of Strategic Priorities (SSP)[23] for Online Safety making clear that Ofcom, in having regard for the SSP, should ensure platforms; *Embed safety by design to deliver safe online experiences for all users but especially children, tackle violence against women and girls, and work towards ensuring that there are no safe havens for illegal content and activity, including fraud, child sexual exploitation and abuse, and illegal disinformation".*

[19] Wood, S. (2026). *Impact of regulation of children's digital lives - Phase II.* Digital Futures for Children centre, LSE & 5Rights Foundation. https://researchonline.lse.ac.uk/id/eprint/138377/
[20] https://www.nuffieldfoundation.org/evidence-and-impact/our-programmes/grown-up/i-love-it-but-i-hate-it
[21] https://www.onlinesafetyact.net/analysis/safety-by-design/
[22] https://www.legislation.gov.uk/ukpga/2023/50/section/1
[23] https://www.gov.uk/government/publications/statement-of-strategic-priorities-for-online-safety/final-statement-of-strategic-priorities-for-online-safety#safety-by-design

Whilst this approach is a core tenet of the OSA, and both Parliament and Government's expectations from it, Ofcom do not define what a safety by design approach must look like in any of their codes of practice. Furthermore, they have not taken a holistic approach to what this means in terms of the design and operation of services, their systems and processes or their business model, even - as we note above - where specific risks relating to features and functionalities have been evidenced in the risk register and corresponding mitigation measures could have already been included in their children's code of practice.[24] While there is broad consensus that safety by design involves proactively building protections into systems and designing out risks to ensure user safety, there remains less clarity around the specific measures platforms must adopt and how these principles translate into operational practice.

The Online Safety Act Network, in collaboration with the 5Rights Foundation, Molly Rose Foundation, NSPCC, End Violence Against Women Coalition (EVAW), Refuge, FlippGen, Glitch and the Internet Watch Foundation (IWF), have developed a Safety by Design Code of Practice which seeks to address this gap and which provides a practical overview of safety by design, set within the framework of the OSA and Ofcom's existing codes and guidance. A PDF of the full code is attached as an annex to this submission.[25]

The Code provides detailed guidance for all tech companies to help them understand safety by design and how safety by design principles might be applied in the context of digital services (including but not limited to services currently within scope of the OSA). It also serves as a template for adoption by the Government and Ofcom as a model for delivering on the Act's requirement, set out in section 1[26], that regulated services are "safe by design" and realising, with no further delay, Parliament's ambitious intent when it passed it.

It is within the Government's gift to expedite this. Firstly, they can mandate that Ofcom produces a Code of Practice, as per the Network's template, as part of their suite of measures to improve children's safety online and make the UK the safest place to be online for all users. This could be achieved by just a few technical updates to the existing legislative framework. Ofcom is already required to produce codes to help services fulfil their safety duties (section 41)[27]. While the Act specifically requires a code dealing with terrorism and one dealing with child sexual abuse material, s 41(3) specifies that Ofcom should prepare one or more codes proposing measures to satisfy the safety duties - and that relates to the illegal content safety duties, the safety duties relating to protection of children as well as the duties on categorised services. Ofcom could already use this power to base its actions, and the safety by design code could underpin the requirements of the other codes. To ensure that this happens and signal their intent, the Government could - with a small targeted amendment - update the OSA to require the production of a safety by design code by Ofcom. Alongside this, the Government should update the OSA to include a definition of safety by design to provide a clear objective for this requirement.

---

24 https://www.ofcom.org.uk/siteassets/resources/documents/online-safety/information-for-industry/illegal-harms/register-of-risks.pdf?v=390983

25 https://www.onlinesafetyact.net/documents/1684/Safety_by_Design_-_Code_of_Practice.pdf

26 https://www.legislation.gov.uk/ukpga/2023/50/section/1

27 https://www.legislation.gov.uk/ukpga/2023/50/section/41

There is a clear political consensus building around a safety by design approach, demonstrated in both the Commons and the Lords during the recent passage of both the Crime and Policing Act and the Children's Wellbeing and Schools Act, which is set against the context of legal action being taken against services in the US for their addictive design. Recent polling has also underlined that this would be popular with the public who believe that this kind of approach - based on product safety testing and risk mitigation - is the least that can be expected[28].

The representative poll of UK adults conducted by YouGov on behalf of the Online Safety Act Network found:

- **84%** are convinced that requiring companies to prove their products are designed and tested to be safe before use would keep all users safe online - the same standard already applied to toys, food, household appliances, and most other products.
- **61%** agree that social media companies take little or no responsibility for designing products that are safe for users.
- **79%** believe we need comprehensive laws to regulate social media platforms because platform operators will otherwise prioritise their business interests over user safety.
- **62%** think that platforms would only take the necessary action if it did not impact their profits.
- **65%** say social media platforms and their leadership, not parents or individual users, should hold primary responsibility for ensuring their products are designed to be safe from the start, followed by the government.
- **Only 2%** think platforms are doing a good job of reducing the risk of harm to users.

In addition

- **75%** believe that AI chatbots must be designed to be fully safe before they can be used.
- **64%** agree platforms must ensure higher levels of safety for children and young people than for adults.
- **43%** feel they have limited or no control over their own safety online

The Government can, and must, use this consultation and the broader momentum behind a more ambitious and considered approach to hold tech platforms, including AI chatbot providers, to account for the harmful design of their products, as is normal practice for other industries. It is only then that the Government will realise what should be its overriding ambition for an online ecosystem which is safe from the start for all users.

28

https://www.onlinesafetyact.net/analysis/uk-public-overwhelmingly-want-social-media-platforms-designed-to-be-safe-before-use-new-polling-reveals/

**Proposed interventions on AI chatbots**

We published a blog on this topic on 21 May.

The Government's approach

As set out above, we believe the terms of the Government's consultation are too narrow, focusing primarily on barring access for under 16s rather than ensuring tech accountability for the harm their products are causing through their design. Yet, the design of a product can impact not just the content outputs, but can also affect behaviour of users, as we discuss below. This narrowness is particularly true of the approach to AI chatbots that has been laid out in the consultation, which offers an extremely limited understanding of the harm caused by chatbots, and fails to properly identify design-based harms - which have been more effectively addressed in the questions relating to social media platforms.

Such discrepancy between approaches is symptomatic of the Government's failure to understand harms related to AI technologies as a continuation and intensification of those already associated with social media and other forms of technology, and which are disproportionately felt by women and girls, Black and minoritised communities, disabled communities and LGBTQ+ communities[29]. Underpinning this is a concern about the datasets on which the LLMs are based and particularly any resulting biases[30] which companies are not appropriately tackling - and, in some instances, seeking to refuse to tackle[31]. This failure has meant that, despite a wealth of evidence demonstrating the harm already experienced by users in the online world, regulation has failed to meet the pace at which AI products such as chatbots have been rolled out onto the market, without proper safeguards or product testing being carried out.

Indeed, whilst AI technologies such as chatbots are posing new and concerning threats to individuals and society, many of the harms connected to chatbots have already been seen and well evidenced in connection to social media platforms. Yet despite this, tech providers have continued to design their products and services without due consideration of the risks posed to individuals. The Online Safety Act Network, alongside 44 leading online safety organisations, have called on the Government to ensure that regulation of AI chatbots mandates proper risk assessments, including mitigation against identified risks, and centres a safety by design approach. They have failed to do so thus far.

The harm

We were pleased that the consultation recognises that "some risks stem from interaction with features rather than content", citing the anthropomorphic qualities of a chatbot that can lead to emotional dependency from users, and, in some cases, the emotional manipulation of users. This understanding of

29 https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes

30 https://dl.acm.org/doi/fullHtml/10.1145/3582269.3615599

31 https://www.courthousenews.com/wp-content/uploads/2026/04/grok-ai-bill-colorado-complaint.pdf

the anthropomorphic features and functionalities as a key driver of harm has been fundamental to the Online Safety Act Network's work and informed our joint regulatory amendment, which is discussed in more detail in the next section.

However, the consultation does not recognise that the features and functionalities deemed risky - such as those that mimic empathy, flattering language, mimicking romantic relationships - are direct consequences of harmful design choices by tech platforms, designed to increase engagement and reliance on these technologies - or in other words, to be addictive - which platforms are actively profiting from. Much like other forms of technology that focus on user engagement, AI chatbots are designed to be addictive, relying on their conversational design and personalisation to encourage continuous engagement. The distinctiveness of this has been recognised elsewhere. In the Chinese legislation (discussed further below) one of the key criteria for triggering regulatory oversight is continuity of interaction[32]. The conversational tone that AI chatbots[33] utilise can give users the illusion of empathy[34], which users may mistake for real emotional connection. Indeed, the hyper-personalised nature of these chatbots mean that young people and adults alike report forming close relationships which blur the boundaries between artificial and real relationships. According to research by Common Sense Media[35], almost a third of American teenagers find chatting to AI companions on platforms like CHAI, Character.AI, Nomi, and Replika, equally or more satisfying than speaking with their friends.

However such relationships can lead to emotional dependency on chatbots, and evidence shows that high usage levels correlates with loneliness, dependence, problematic use and lower socialisation[36]. For users who have formed this dependency, the impact on their mental health when AI chatbots are no longer available to users who have formed emotional attachments with chatbots[37], either due to changes in design or by the introduction of paywalls, may be detrimental, with some experiencing withdrawal symptoms similar to those seen in substance addiction[38]. Instead of seeking to age-gate these functions which are demonstrably harmful, we need proper regulation that seeks to address those design choices and ensure that no unsafe product is rolled out onto the market without rigorous risk assessments, product testing and mitigation - otherwise children and all users will continue to face harm from design choices by providers made to keep users hooked on their products.

---

[32] https://www.hoganlovells.com/en/publications/chinas-interim-measures-for-the-administration-of-anthropomorphic-ai-interaction-services

[33] https://thealiendesign.medium.com/enhancing-customer-engagement-with-ai-powered-chatbots-the-key-to-seamless-interactions-91783ed0a14a

[34] https://www.researchgate.net/publication/385980288_The_Illusion_of_Empathy_How_AI_Chatbots_Shape_Conversation_Perception/link/673ea37bc1b80e5616501f84/download?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19

[35] https://www.commonsensemedia.org/sites/default/files/research/report/talk-trust-and-trade-offs_2025_web.pdf

[36] https://www.media.mit.edu/publications/how-ai-and-human-behaviors-shape-psychosocial-effects-of-chatbot-use-a-longitudinal-controlled-study/

[37] https://www.bmj.com/content/391/bmj.r2239

[38] https://lego17440.medium.com/the-male-loneliness-crisis-and-the-rise-of-ai-companions-a-digital-band-aid-or-a-path-forward-b8215b93eb7f

Whilst we were also pleased to see "hallucinations or false/ misleading responses" included in the list of potentially risky features for young people, it is not clear why false information - or disinformation - would be considered harmful to young people but permissible for adult users. There is already extensive research into the adverse impact that chatbots are having on the information environment. Outdated training data has the potential to contribute to false or harmful information being spread on anything from false healthcare advice about the link between vaccines and autism[39], advice to put glue on to of your pizza by Google's AI overview[40], and incorrect information about telescopes on Google's AI chatbot Bard which wiped $100bn off its share price[41]. Indeed one study found that whilst AI chatbots will often access the correct information to provide an answer, the interpretation of that information will be incorrect or false, a phenomenon that they name 'certain hallucinations overriding known evidence'[42]. While citations have been suggested as a mechanism to show users the provenance of information and to allow them to perform further checks and research, LLMs do make up citations - this is a sub-set of chatbots' tendency to "hallucinate". AI chatbots have also been reported to be 'overly flattering or agreeable'[43], following research that demonstrated sycophancy in large language models (LLMs)[44]. Sycophancy can be dangerous when advice and support is provided by a chatbot that reinforce harmful implicit assumptions, beliefs of actions in order to satisfy the needs of users, contributing to the spread of false information. The risk associated with false information, and the risky way in which the personalised tone of a chatbot can further encourage trust in false information, is harmful to both adults and children alike, and should be recognised as a fault that should have been designed-out at the point of development.

This concern extends to all the features and functionalities set out in the Government's consultation that are deemed 'dangerous', particularly for vulnerable adult users. Our research brief[45] outlines countless reports that speak to the addictive nature of these technologies, and the detrimental impact they can have on adults' mental health and wellbeing, the normalisation of both misogynistic and racist attitudes, the pollution of the information ecosystem as a result of false information. These are not harms that only impact children, and they should not be allowed to occur in the first place. Indeed, whilst the consultation refers to the research relating to the harmful impact of chatbots as "emerging", we argue that there is already a significant body of evidence that attests to the harm being experienced by individuals and the impact on society. We should not wait for more users to be harmed to provide us with the right sample size.

Despite the Government's current framing in the consultation that AI can be used to support disadvantaged students in their learning, there are concerns about the use of AI and the long-term

[39] https://www.acpjournals.org/doi/10.7326/ANNALS-24-03933
[40] https://www.bbc.co.uk/news/articles/cd11gzejgz4o
[41] https://www.bbc.co.uk/news/business-64576225
[42] https://arxiv.org/html/2502.12964v2
[43] https://openai.com/index/sycophancy-in-gpt-4o/
[44] https://arxiv.org/html/2505.13995v1
[45] https://www.onlinesafetyact.net/analysis/ai-chatbots-the-case-for-action/

impact on learning[4647]. UNESCO's guidance on AI and education points to worsening digital poverty, AI that is outpacing national regulatory adaption, reduction of diverse opinions and furthering marginalising already marginalised voices, AI-generated content polluting the internet and a lack of understanding of the real work as some of the long-term impacts of using AI in education[48]. Given that, whilst in its nascency, there is already evidence of harm in using AI products in education, the Government must apply the precautionary principle and ensure this technology is well understood and product tested before entrenching these systems in our classrooms. Yet despite this, and despite the fact that there is "no clear guidance yet about how to measure the impact of AI in education to show whether it is effective"[49], the Government has concluded that "the biggest risk would be doing nothing"[50], and that our priority must be rolling out this technology quickly[51].

Regulatory landscape

The Government's narrow amendment[52] relating to AI chatbots and illegal content does not address harm resulting from the anthropomorphic features and functionalities and other design-based harms that have already been outlined. The Government's decision to include these design-based harms in the consultation is a clear admission of the regulatory gap that remains. Indeed, just as harms arising from AI chatbots are not just about content, nor are they just about children's access. They are about unsafe, untested products - designed to be addictive and manipulative - being rolled out without regulatory oversight. Such a regulatory approach is possible: for example, it has recently been introduced in China. The Chinese regime "applies to products or services that utilize AI technology to provide the public within the territory of the People's Republic of China with simulated human personality traits, thinking patterns, and communication styles, and engage in emotional interaction with humans through text, images, audio, video, etc"[53]. In particular, "security assessments are required", covering a wide range of issues but including "governance of training data, mechanisms for identifying and responding to extreme user scenarios, user demographics and behavioral patterns, protection measures for vulnerable groups

[46] https://www.404media.co/microsoft-study-finds-ai-makes-human-cognition-atrophied-and-unprepared-3/
[47] https://link.springer.com/article/10.1186/s41239-024-00444-7
[48] https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research
[49] https://www.gov.uk/government/publications/ai-in-schools-and-further-education-findings-from-early-adopters/the-biggest-risk-is-doing-nothing-insights-from-early-adopters-of-artificial-intelligence-in-schools-and-further-education-colleges
[50] https://www.gov.uk/government/publications/ai-in-schools-and-further-education-findings-from-early-adopters/the-biggest-risk-is-doing-nothing-insights-from-early-adopters-of-artificial-intelligence-in-schools-and-further-education-colleges
[51] https://www.gov.uk/government/publications/generative-artificial-intelligence-in-education/generative-artificial-intelligence-ai-in-education
[52] https://www.legislation.gov.uk/ukpga/2026/20/section/248/enacted
[53] https://www.cac.gov.cn/2025-12/27/c_1768571207311996.htm

such as minors and the elderly, complaint-handling processes, and remediation of identified risks"[54]. The rules also require special protections for minors[55].

We were extremely disappointed that the Government failed to use the opportunity presented to them through the Crime and Policing Act to introduce much more comprehensive measures to prevent this. The Government had an opportunity during the Crime and Policing Act to accept an amendment from Baroness Kidron to extend the scope of their proposal and make it an offence for chatbot developers not to carry out risk assessments for a range of issues, including the creation of illegal content and harmful design choices[56]. They refused this.

While it is true that that the legislative vehicle constrained the scope of the Government's options and Kidron's approach - using the criminal law to address a regulatory cap - was not ideal, the Government also had rejected an opportunity previously - via the Product Regulation and Metrology Act - to bring chatbots and other AI-enabled digital products into that legislation and make them subject to product safety standards (the PRAMA only covers physical products with an AI capability). It has also failed to bring in an AI Bill, with no new legislation promised in this year's King's Speech.

This despite the fact that stronger regulation of AI chatbot providers is supported by the public, with recent polling carried out by the Online Safety Act Network found that 75% of the UK public believe that AI chatbots must be designed to be fully safe before they can be used[57]. Indeed this would be far more effective, given that evidence published on the 21st of May by Ofcom demonstrates that age-assurance is not currently keeping young people safe on major platforms - with children as young as 11 still regularly accessing social media and harmful content still proliferating unchecked[58]. By proposing age-gating of certain features as the proposed solution to design-based harms, the Government now risks leaving vast numbers of UK citizens without proper protections.

Furthermore, as highlighted in further detail above, the Government's proposed approach to children's safety online - age-gating over tech accountability - is at direct odds with the growing role of Ed Tech within the Department for Education, particularly the use of AI Tutors which they have referenced themselves[59]. Rolling out these technologies for "disadvantaged pupils" - vulnerable communities -

---

54 https://www.hoganlovells.com/en/publications/chinas-interim-measures-for-the-administration-of-anthropomorphic-ai-interaction-services

55 https://carnegieendowment.org/research/2026/02/china-is-worried-about-ai-companions-heres-what-its-doing-about-them

56 https://www.onlinesafetyact.net/analysis/the-uk-must-hold-ai-chatbot-companies-to-account-for-the-harm-they-cause/

57 https://www.onlinesafetyact.net/analysis/uk-public-overwhelmingly-want-social-media-platforms-designed-to-be-safe-before-use-new-polling-reveals/

58 https://www.ofcom.org.uk/siteassets/resources/documents/online-safety/protecting-children/project-mercury/update_tech-firms-responses-to-our-call-for-action-to-protect-children.pdf?v=418243

59 https://roadmap-for-modern-digital-government.campaign.gov.uk/ai/ai-in-education/

without proper regulation risks further entrenching inequality within the education system[60]. There is therefore a profound lack of clarity about how departments across Government are working together on this issue, including those across DSIT, the Home Office, and the Department of Education. A robust regulatory approach would mean that there is trust and confidence in the safety and security of AI products being used in the home as well as those being used in the classroom - as all products will be required to have been rigorously tested before they hit the market.

Yet despite this, the Government has no plans to bring forward an AI Bill in the forthcoming legislative programme or further online safety legislation. This will lead to a repeat of the chaotic, piecemeal approach seen at the end of the previous Parliamentary session where the Government was forced to respond to pressure from Parliamentarians and campaigners with a series of individual amendments to different Bills - notably the Crime and Policing Bill (now Act) and the Children's Wellbeing and Schools Bill (now Act) - which failed to fully address the scale of the harm related to AI. This lack of legislative vehicles to address significant gaps in the online safety regime is potentially compounded by the inclusion of a "Regulating for Growth" Bill which risks removing vital existing protections for consumers (data protection, consumer protections, online safety) in favour of boosting innovation and economic growth.

Safety by design

Over the last six months, the Online Safety Act Network has been coordinating a working group of civil society organisations working on a cross-harm basis to develop policy solutions to the harm arising from AI chatbots. The group proposed a holistic approach to the regulation of AI chatbots to ensure no unsafe product hits the market - one that requires proper governance, risk assessment, product testing and mitigation throughout the entire lifecycle of the product - and that recognises the anthropomorphic features and functionalities of an AI chatbot as a unique driver of harm[61]. In collaboration with this group, we developed an amendment to the Online Safety Act that centres this approach and ensures that risk is both assessed and mitigated against, and that products are rigorously tested. This amendment is provided in full as a PDF attachment.[62]

With glaring gaps still in need of attention, the Government must commit to ensuring their regulatory approach is one that centres the principles of proper risk assessment, product testing and safety by design, or risk even more preventable harm related to chatbots. They will need to bring forward new legislation to do so - but despite the clear demand for further action there was no sign from the King's

60 https://www.gov.uk/government/news/edtech-and-ai-companies-invited-to-help-build-safe-ai-tutoring-tools-for-disadvantaged-pupils

61 https://www.onlinesafetyact.net/analysis/the-uk-must-hold-ai-chatbot-companies-to-account-for-the-harm-they-cause/

62 https://www.onlinesafetyact.net/documents/1717/AMENDMENT_TO_THE_ONLINE_SAFETY_ACT_ON_AI_CHATBOTS.pdf

speech that they will bring forward new legislation to tackle the growing threat of AI or strengthen the OSA. We have been clear that age-gating particular features and functionalities alone will not address the systemic harm caused by these products to all users, nor will it provide a robust barrier to children experiencing harm, who have demonstrated that they can easily circumvent these 'protections'.

**Wider interventions to strengthen the Online Safety Act and deliver sustainable change**

As above, we are fully aware of the context in which the Government was forced to launch this consultation and the imperative behind the narrow framing of it. But we are strongly of the view that, if this is the only chance to make changes to the regulatory landscape for the foreseeable future, the Government must do all it can to make its package of proposals as strong and comprehensive as possible and also to signal that policy development work will be undertaken to address some of wider, systemic challenges that arise from the rapid development of online technologies.

With that in mind, we draw the Government's attention to the following additional proposals:

- The Children's Coalition on Online Safety's statement, which we have co-signed[63].
- The forthcoming statement of policy recommendations from the IWF and other child protection partners, which also endorses our 10 Point Plan for Strengthening the Online Safety Act.
- The recommendations in the Ada Lovelace Institute Report on children's digital lives, which endorse the adoption of our safety by design code of practice along with the need for action on AI chatbots and wider AI regulation.[64]
- The recommendations from Internet Matters in their recent report on the Online Safety Act, which include action on safety by design[65]
- [Something from EVAW?]

More broadly, we urge the Government to invest policymaking resource into the following areas to ensure that regulation keeps up with emerging evidence of harm.

- **Specific powers to short circuit lengthy processes when there is an urgent high harm risk from individual platforms** (such as suicide sites) and to compel immediate takedown of content related to priority offences. The legislation needs to be clear that Ofcom can use its power 'to direct' companies rapidly in an emergency - for example, with regard to deprioritisation, account suspension as well as takedown orders - with safeguards such as a company's right to appeal against those measures or a requirement on Ofcom to apply to the Court (for emergency powers and business disruption). In addition, specific crisis rules and protocols need to be introduced to compel action at times of public emergency.

63 https://5rightsfoundation.com/resource/statement-on-the-uk-governments-growing-up-in-the-online-world-consultation/

64 https://www.nuffieldfoundation.org/evidence-and-impact/our-programmes/grown-up/i-love-it-but-i-hate-it

65 https://www.internetmatters.org/hub/research/online-safety-act-report-2026/

- **Action on app stores:** this is unfinished business from the OSA; there is a requirement on Ofcom to produce a report but this is currently not expected until early 2027. Legislation is needed to bring in a new statutory Code of Practice for app store and device operating systems, without waiting for the publication of the report. This should be seen as a complement to, not a substitute for, effective age assurance at platform level.
- **Risk assessment and mitigation obligations** should be placed on providers of products containing AI systems (including products aimed at children), and the Product Regulation and Metrology Act should be updated to cover digital as well as physical products.
- **Protection for citizens (including children) from the AI-exacerbated spread of misinformation**, as per the recommendation from the SIT committee report last year, with generative AI platforms brought into line with other online services that pose a high risk of producing or spreading illegal or harmful content.
- **Action to address societal and information ecosystem threats:** bringing in foreign ownership rules for social media companies and other parts of the digital information ecosystem and a fit and proper ownership test; reviewing the role of public service broadcasting, and due prominence of quality news; putting the National Security Online Information Team on a statutory footing.
- **Preventing harms arising from the profit-driven business model:** this might include a conduct-based set of measures to address the behaviours and incentives that apply to tech managers and shape the commercial and product strategy of regulated firms i.e. as we see in financial services; and the introduction of "Know your Customer" requirements - as per financial services - to prevent the use of anonymous accounts or bot networks to cause individual and societal harm, or to carry out fraud.
- **Action on online advertising** - successive governments have kicked this into the long grass, the Online Advertising Taskforce has led to no meaningful change and the Online Advertising Programme has never been implemented. This is a significant omission given the role of the advertising-based business model to fuelling online harms.
- **A stronger framework for regulatory cooperation and enforcement** across the digital and online sphere that is fit for the challenges ahead.

**Online Safety Act Network**
**May 2026**