



STRENGTHENING THE ONLINE SAFETY ACT: A TEN-POINT PLAN FOR GOVERNMENT

The Online Safety Act Network works on a cross-harm basis and this plan is supported by 23 organisations spanning interests including child protection, suicide and self-harm, online hate and violence against women and girls. The OSA is a significant milestone but the pace of change is too slow and the gaps arising from Ofcom’s approach to implementation are clear. We call on the Government to adopt this plan as a matter of urgency to ensure the delivery of the Act’s online safety objectives.

We urge the Government to make the following amendments to the Online Safety Act 2023

Risk assessment and mitigation

- Strengthen the obligation in the Act that regulated services take reasonable steps to seek to **reduce the risk of harm** to users as identified in their OSA risk assessments.
- Revise the “clear and detailed” and “technically feasible” conditions for code of practice measures to allow for **more stretching, outcome-focused measures in the OSA codes**.
- **Remove the safe harbour** provision.

Safety by design

- Insert a definition of **safety by design** into the Act to make clear to Ofcom and services what Parliament intended, and lay secondary legislation to require Ofcom to produce a “safety by design” code of practice, a cross-cutting code to sit beneath the existing codes. This should also clarify the requirement that services address harms to children caused by addictive or compulsive design.

User protections

- Make clear that, once illegal content or imagery has been removed, further shares of the same material are removed without further moderation requirements: **a ‘stay-down’ provision**.
- Deliver Parliament’s intent that services can be categorised based on size OR risk and bring **small but risky platforms** into category 1
- Set a **minimum standard for Terms of Service (ToS)** for category 1 platforms and include a **“no rolling back” requirement** such that ToS and safety measures must be maintained to an equivalent or greater protection to that at the time of OSA Royal Assent.

Systemic issues

- Insert an **overarching duty of care** into the legislative framework to require a more systemic approach to risk assessment and mitigation from regulated services with regard to regulated content.
- Impose a **duty on Ofcom to cooperate with other domestic regulators** and put the Digital Regulation Cooperation Forum on a statutory footing.
- Insert a new series of standalone **duties for AI chatbot services** to address the emerging harms that currently fall out of scope of the OSA.

Further detail on these amendments is provided in the attached briefing

THIS PLAN IS SUPPORTED BY



ANTISEMITISM
POLICY TRUST



Plus, **Adele Zeynep Walton**, campaigner with Families and Survivors Against Online Suicide Harms

STRENGTHENING THE ONLINE SAFETY ACT: A TEN-POINT PLAN FOR GOVERNMENT

Issue

The Online Safety Act is long and complex but Parliament's intent when it passed was clear. It is set out in the very first section of the Act: "making the use of internet services regulated by this Act safer for individuals in the United Kingdom" with the duties imposed on providers intended to secure "among other things" that services regulated by the Act are "safe by design" and "operated in such a way that a higher standard of protection is provided for children than for adults".

Two years' on from the Act's Royal Assent, with the two main safety duties now in force, it is unlikely that Parliament's expectations will be delivered any time soon. As documented extensively by the OSA Network and many of our partners, Ofcom's narrow interpretation of a number of key provisions in the Act has led to an overly cautious, risk-averse approach to drawing up the codes of practice for the two main safety duties, on illegal harms and protection of children. The regulator's chosen interpretation of the Act does not use the full scope of the available provisions to deliver the outcomes that Parliament envisaged and, crucially, does not reflect the intention that services be safe by design: an intention the former Secretary of State further emphasised in his [Statement of Strategic Priorities for Online Safety](#). The result is a gap between the risk assessment requirements and the mitigation obligations, the narrow, limited measures in the codes of practice, and a suite of unambitious ex-post measures to reduce the impact of harm that has already occurred rather than upstream, content-neutral, "by-design" interventions to seek to prevent it occurring in the first place.

The current Government has been forced repeatedly to justify Ofcom's approach, pleading with critics to "wait and see" while the Act is implemented. But these issues will not be resolved by maintaining the same interpretative approach. It is squarely within the Government's gift to solve this problem: it can amend the Act, through a set of targeted amendments to remove the barriers that Ofcom believes it is facing and bring the implementation of the regime closer to the Act's online safety objectives that had cross-party agreement during the Bill's passage. Failure to act will leave these shortcomings baked into the regime for years to come, regardless of how many iterations of the codes or additional measures the regulator brings forward.

We believe that such action by the Government would have broad public and cross-party support. Parliamentarians in both Houses have become increasingly frustrated with Ofcom's approach to implementation (see [House of Lords Motion of Regret](#) 30/10/2025; [SIT Committee Report](#) July 2025.) Recent polling by [More in Common](#) found that amongst UK parents of all political persuasion, 63% thought that politicians aren't taking the issue of keeping children safe online seriously enough and 67% would feel more positively towards a political party that supported increasing online safety protections.

We group our 10 proposed amendments below, demonstrating how they will work together within the framework of the legislation to get its implementation back on track. We will be publishing text for these soon. We urge the Government to bring these forward as soon as possible so that no further time is lost.

The solutions

Risk assessment and mitigation

These are small, technical amendments to the Act which do not affect the overall regulatory framework but which, taken together, will remove the barriers Ofcom believes prevent it taking a more ambitious approach to harm reduction. This is not a matter of interpretation. Without these amendments, Ofcom will not deliver on the online safety objectives enshrined in the OSA at [Schedule 4, Section 3](#).

These three proposals are interlinked.

Risks: Ofcom has provided extensive evidence in its risk registers on the risks of exposure to content covered by the Act related to different features and functionalities. The codes of practice provide only minimal measures to address these features and functionalities: the gap is evidenced [in this table](#). Ofcom could have proposed that services take steps to address all the risks arising from their risk assessment but chose not to; they argue that this is because the Act requires the measures in codes to be “clear and detailed”. We do not agree with this interpretation the Act in each of the safety duties imposes an obligation to mitigate against risks found in the risk assessment yet Ofcom’s approach to interpretation has undercut these clear statutory obligations. The previous Government was also clear, when taking the Bill through its Lords’ stages, that outcome-defined obligations for regulated services were appropriate in such a regulatory regime: the Bill Minister, Lord Parkinson, told the Lords on [25 April 2023](#) that “Ofcom’s guidance and codes of practice will set out how they can comply with their duties, in a way that I hope is even clearer than the Explanatory Notes to the Bill, but certainly allowing for companies to have a conversation and ask for areas of clarification, if that is still needed”.

Code measures: Limitations on the measures in the codes do stem directly from the Act; [Schedule 4 2 \(b\)](#) sets out that measures “must be sufficiently clear, and at a sufficiently detailed level, that providers understand what those measures entail in practice”. It also requires them to be “technically feasible” and “proportionate”. Ofcom’s interpretation means that - instead of placing the responsibility on companies to mitigate the risks on their service in a way that is appropriate (and proportionate and technically feasible) for them - the regulator has taken upon itself to define those measures, and only when there is evidence that those measures work - eg they are already in operation on some services. So, the measures are set at current practice levels, rendering the “online safety objectives” that Parliament set out, very clearly, in [section 4 of Schedule 4](#), redundant. We also have concerns about the impact of the “technically feasible” condition, combined with the “proportionality” condition, to be used as a get-out clause for companies. See our paper [here](#).

Safe harbour: Ofcom’s approach is further compounded by something that is beyond their control: the inclusion in the OSA of the “safe harbour” provision (at [s 49 \(1\)](#)), so that compliance with codes means compliance with safety duties. There is no incentive to go further than the measures in the codes, even if companies know that these do not address all the risks on their services, or are less than they currently do. Nor is there any incentive to innovate: to develop new and different measures to address known risks, from which Ofcom could then draw for future code measures - raising the bar across the sector.

The codes have to be amended: the regulator has taken on all the burden for sector-wide safety improvements on itself. There is also nothing to stop services that already have additional safety measures in place removing them and remaining in compliance with their duties (see point 4, below).

It is incumbent on the Government therefore to amend the Act to shift Ofcom’s implementation of the regime towards a more risk-based, outcome-focused approach that puts the onus on services, rather than the regulator, to keep users safe and delivers the legislation’s online safety objectives.

We propose the following amendments:

- Strengthen the obligation in the Act that regulated services take reasonable steps to seek to reduce the risk of harm to users as identified in their OSA risk assessments by amending Schedule 4.
- Revise the “clear and detailed” and “technically feasible” conditions for code of practice measures to allow for more stretching, outcome-focused measures in the OSA codes.
- Remove the safe harbour provision.

Safety by design

Although Ofcom’s more recent work does include express reference to safety by design, its interpretation of “safe by design” is limited mainly to a few ex-post measures (such as their proposals on recommender algorithms); 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 where indicated by the risk register. We have [written extensively on this](#). If Ofcom continues its measure-by-measure approach to iterating the codes of practice, there will be no change in the profit-driven, attention-grabbing design choices by companies; the regulator’s interventions will always be years behind the harm and the promise of the OSA will be unfulfilled.

The Government took steps to set out to Ofcom the importance of safety by design in its Statement of Strategic Priorities. Ofcom’s response to this - a restatement of what they are already doing - shows how little they understand this objective and its importance to delivering improved online safety. The Government must therefore act more robustly to ensure that the regulator takes this seriously by clarifying the obligations that are already in the OSA. We propose the following amendments:

- Insert a definition of safety by design into the Act to make clear to Ofcom and services what it was that Parliament intended, and lay secondary legislation to require Ofcom to produce a “safety by design” code of practice, a cross-cutting code to sit beneath the content-based codes. This should also address issues relating to harms to children caused by addictive or compulsive design, which are referenced in the children’s risk assessment duties ([section 11 \(6\) \(f\)](#)) but which have not been addressed by Ofcom in their work to date.

User protections

There are a number of interventions which the Government needs to make to ensure that the baseline of user protections envisaged by Parliament when it passed the Act is delivered. These include a clarification that content identified as illegal content and removed should stay down from then on, such that subsequent shares do not require the same moderation process to be followed; this an issue which has become particularly pertinent in relation to [non-consensual intimate image abuse](#). Although Ofcom's Illegal Content Judgment Guidance in relation to NCII now suggests that providers can presume that repostings of an intimate image found to be non-consensual are also non-consensual, this is a problem that potentially affects many types of illegal content. We endorse the calls to strengthen the VAWG guidance - [A Safer Life Online for Women and Girls](#) - to become a code of practice to ensure Ofcom have the power to enforce against companies that fail to comply. The [problem created for the Government by Ofcom's advice](#) to the previous Secretary of State on categorisation needs to be resolved to address the increasing evidence emerging of the role of small platforms in the spread of extremism, radicalisation, [far-right](#) and misogynistic ideologies. Finally, there is no baseline measure for protections for users under platforms' terms of service within the Act; indeed existing safety measures can easily be rolled back (as we saw with Meta's announcements earlier this year); companies that do more than the measures in the codes, do not have to continue to do so (as we explain above). See our commentary [here](#).

The Government's wider strategic objectives relating to countering extremism, community cohesion and violence against women and girls will not be delivered without action to tighten up the Online Safety Act's existing duties in relation to illegal content, small but risky platforms and user protections. We propose the following amendments:

- Make clear that, once illegal content or imagery has been removed, further shares of the same material are similarly removed without the need for further moderation process: a 'stay-down' provision.
- Deliver Parliament's intent that services can be categorised based on size OR risk by amending [Schedule 11](#) to bring small but risky platforms into category 1.
- Set a minimum standard for Terms of Service for category 1 platforms, mirroring the criteria set out in the user empowerment duties ([section 15](#)), and include a "no rolling back" requirement such that ToS and safety measures must be maintained to an equivalent or greater protection to that at the time of OSA Royal Assent.

Systemic issues and gaps

Our final set of proposals relate to more substantive Government interventions to address some of the gaps that are now apparent in the Act itself, rather than as a result of Ofcom's implementation of it. The first - the inclusion of an overarching duty of care - is a medium-term proposal which, we believe, would future-proof the Act and overcome the limitations of the iterative approach to codes of practice taken to date by Ofcom. The second - a duty on Ofcom to co-operate with domestic regulators - was rejected during the passage of the Bill as there was an assumption that Ofcom would do this by virtue of its

existing operational arrangements; emerging concerns, for example, relating to the intersection of age verification with privacy and data protection regulations suggests that the relationship between Ofcom and the ICO is weak, at best, despite the Digital Regulators Cooperation Forum. Similarly, a silo-based approach to digital regulation will not address the challenges of cross-cutting technological advances, such as AI. Related to this, the third issue - the partial coverage of AI chatbots by the existing legislation - is in urgent need of resolution (as our work here (tbc) sets out).

We propose the following amendments:

- Insert an overarching duty of care into the legislative framework to require a more systemic approach to risk assessment and mitigation from regulated services, including on emerging harms arising from new products and services.
- Impose a duty on Ofcom to cooperate with other domestic regulators (to mirror the duty regarding overseas regulators) and put the Digital Regulation Cooperation Forum on a statutory footing.
- Insert a new series of standalone duties (as per the duties on pornography providers in section 5) for AI chatbot services to address the emerging harms that currently fall out of scope of the OSA.

Online Safety Act Network
January 2026
hello@onlinesafetyact.net

Text for the 10 proposed amendments will be available in due course.