



A roadmap for a better online future

A new online safety settlement
for children, parents and families



Contents

Summary	3
Our five-point plan to transform children’s safety and wellbeing online	10
1. Fixing and decisively strengthening the Online Safety Act	11
2. Extending the Online Safety Act to cover Wellbeing-by-Design	17
3. Requiring transparency, accountability and candour from Big Tech	21
4. A ‘polluter pays’ and whole stack approach to harm reduction	24
5. Education as inoculation – critical digital and media literacy that protects young people from harm and prepares them for our future economy	27

Summary

For children's online safety and wellbeing, this year marks an inflection point.

It's clear that decisive and radical action is needed to turn the tide on the preventable harm that costs young lives and blights a generation.

This is our urgent but hopeful roadmap for change.

Progress has undeniably been made. Fledgling regulation is in place. In the UK, calls for decisive intervention have never been stronger.

However, the risks are also growing. Geopolitical headwinds have stalled progress, and the flaws of an imperfect regulatory regime have been highlighted by Ofcom's deeply unambitious approach to enforcement.

Parents are clamouring for meaningful action. After years of delays and insufficient action, they understandably feel let down by successive governments, regulators and most of all by tech firms that have consistently prioritised profit over children's safety.

2026 is the year we must act – decisively, boldly, and with the courage to deliver meaningful and comprehensive change.

Our roadmap for a better online future

This report sets out a roadmap to deliver the meaningful change that children and families demand and deserve. We outline five major steps that this Government should take to decisively protect children and young people from preventable harm – and to demonstrate to parents that decisive action is on the way.

This is a bold and comprehensive plan that, if backed by political will, can attract the support of experts, civil society, young people and a clear majority of parents. **Three-quarters (73%) of UK adults support new legislation to strengthen regulation and better protect children and young people from harm,¹ with support for regulation stronger than recent support for social media bans.²**

Among those who support fresh legislation, an overwhelming majority (84%) want to see it introduced this year.³

It is clearer than ever that preventable online harm is first and foremost an issue of product safety. As in every other part of the economy, the answer to market failure is strong and effective regulation.

The Online Safety Act is therefore a crucial part of the solution. However, after years of delays and watering down, its implementation has fallen badly short. Ofcom oversees a triangulated regime that fails to tackle the incentives and business models that perpetuate harm and continue to cost young lives.

A strong and ambitious regulatory fix is now required – part of a broader set of interventions across the tech and policy stack, marking an inflection point in our response.

The Government must commit to regulation that is more than a tick-box exercise, more than a sticking plaster approach. **Labour committed to strengthening the Act while in opposition. Parents will now rightly expect this promise to be kept.⁴**

1 Research conducted by Savanta for Molly Rose Foundation. 2,048 UK adults were surveyed in January 2026.

2 Polling undertaken the Good Growth Foundation found that 66% of UK adults support a legal ban on social media for children under 16.

3 Savanta polling for MRF.

4 Labour's now Deputy Leader Lucy Powell announced in January 2023 that an incoming Labour Government would introduce stronger legislation as a 'top priority', with a focus on algorithmically recommended harmful content. She told The Observer: "I met many of the families who have lost teenagers from online activity, and I promised them we would act. We owe it to all those who have been harmed online to get this right." <https://www.theguardian.com/technology/2023/jan/01/labour-pledges-toughen-online-safety-bill>

A five point settlement to save lives and tackle online harm

In this report, we set out five key measures that the Government should take, and that collectively represent a bold new online safety settlement for the UK's children and parents.

As a matter of urgency, the Government should:

1. Fix and decisively strengthen the Online Safety Act

Regulation remains the most powerful tool we have to decisively turn the tide on preventable online harm, and to intervene in the incentives and business models that continue to cause harm on social media, gaming platforms, messaging sites and AI chatbots.

However, regulation is not working as intended. The Government must move quickly to fix the foundations, with a series of legislative amendments that can immediately address the most pressing constraints in the regulatory regime.

The Government should then go further – introducing a new Act in the next Parliamentary session that builds upon and radically strengthens the existing approach.

It is time for strengthened approach to regulation, with the introduction of a systemic, outcome-focused Duty of Care, and new conduct-based measures that can successfully build on lessons learned from the regulation of financial services. These measures will collectively start to bite on the incentives and operating culture of some of the largest and most cash-rich corporates in the world.

The Government must also ensure minimum age limits are robustly enforced. Ministers should proceed with a new set of strengthened risk-based age ratings that see minimum joining ages determined on an assessment of age-appropriateness and risk, much like film age ratings, thereby incentivising the market to adopt lower-risk functionalities if they wish to serve younger groups. The highest risk functionalities, including livestreaming and AI chatbots with human personas, should carry the highest age ratings.

2. Extend the Act to cover wellbeing, and make wellbeing-by-design the price of admission to the UK market

The Government should commit to broadening the scope and ambition of the Online Safety Act – widening its objectives to not only necessitate harm reduction but to actively promote and protect children's wellbeing by design.

Through new duties on platforms, and new strategic objectives on Ofcom, we can introduce a robust and demanding set of measures that make it clear that children's wellbeing is the price of admission to the UK market.

We need a bold and ambitious set of measures that mark an end to addictive and harmful design choices, and a new duty to ensure digital products are built to be age-appropriate, high quality and nourishing by design.

Algorithms should be fundamentally repurposed – if a platform intends to use them, they should not only be free of harmful and toxic material, but should nourish and support our children’s development, prioritising quality, diversity and positive value.

Feeds should be made to recommend high-quality, age-appropriate content from a diverse range of trusted sources, including trusted mental health support, education providers and public service broadcasters.

3. Require new levels of transparency, accountability and candour from Big Tech

Transparency is a powerful lever for change – and one of the most powerful tools in our arsenal to mitigate harm and shape digital markets that can work for children, families and society. However, as it stands, tech companies continue to operate with relative opacity, a lack of transparency that impedes regulation and underpins extractive business models.

A fundamental reset is required. This means new duties mirroring those in financial services that require the largest tech firms – and their senior decision makers – to proactively disclose information about which the regulator would expect to be made aware.

We need a new approach to transparency across the stack, with a new deal for advertisers who unwittingly monetise social media harm, and whose advertising spend – at the core of Big Tech’s business model – is attritionally lost to their opaque advertising systems and models.

The Government should also promote new accountancy standards that can draw on lessons from tackling climate change, requiring corporates to report on their exposure to online safety risks, both directly and across their supply chain.

4. Adopt a ‘polluter pays’ and whole stack approach to harm reduction

We need a new whole-system approach to ensure that regulation works, and to meaningfully drive down children’s exposure to harm.

This can be delivered through the adoption of a ‘polluter pays’ funding approach, with the extension of the industry levy that currently funds Ofcom to also support civil society and academia, and to pump-prime the production of research and data that can actively inform and support the regulator’s investigatory and enforcement work.

This also means a new commitment to tackling harm across the stack – bringing app stores into the scope of regulation, and introducing a new Code of Practice that sets minimum standards for interventions at device level and operating systems, including new more robust parental controls.

A new Act should include a legislative commitment to require the use of technically feasible device-level measures to protect children from being exposed to grooming and abuse, principally through the default use of nudity detection software on children’s devices.

Coerced self-generated images are not only appalling in their own right, but are also being used by group-based offenders (known as Com groups) to perpetuate a broader range of sadistically motivated crimes, including coerced acts of self-harm, child sexual abuse and even suicide.⁵

While the Government's recent commitment to make it 'impossible' for a child to take, send or share a naked image is welcome, a new Act should make it a legislative requirement rather than a choice for the likes of Apple and Google to disrupt and prevent child abuse and grooming at source.

5. Prioritise education as inoculation, with a new strategic focus on critical digital and media literacy

It's time for a bold reset of school-age education, starting with strategic recognition of the importance of critical digital and media literacy education. High-quality, cross-cutting digital education can not only inoculate children against the worst effects of online harm, it can equip them with the skills they need to flourish in the AI and digital economy of tomorrow.

Building on England's Curriculum Review, we should invest in foundational skills around platform and algorithmic literacy, giving young people a critical understanding of how digital environments work and shape our lives – including the role of algorithms, data and persuasive design.

The Government must also support schools to robustly embed critical digital and media literacy across the entire curriculum and all age groups, while training every teacher to take responsibility for building children's digital life skills.

Together with stronger regulation, high-quality digital and media literacy education can help to turn the tide on online harm, draining the power from algorithms that continue to expose young people to appalling harm.

At the same time, we can meaningfully invest in the brain capital and economic potential of a generation, giving young people the critical thinking, digital citizenship and practical skills they need to thrive in an AI economy, prepare for a new voting age of 16, and deal with the increasing threats a fractured information ecosystem poses to our democracy and national security.

However, we can't realistically expect to give young people the skills they need if we simultaneously prevent them from using the very platforms we want them to be able to critically reflect on.

⁵ Resolver Trust and Safety, in partnership with Molly Rose Foundation (2026) Weaponised Loneliness: Critical Harm Intelligence Briefing.

Why an Australia-style social media ban is not the right approach

In recent weeks, calls for a social media ban for under 16s have grown considerably. These calls are a symptom of profound concern around children's online safety, and parents are right to demand bold and comprehensive further action.

At Molly Rose Foundation, we strongly believe that bans are the wrong approach. Though well-intentioned, they risk doing more harm than good – causing harms to migrate to other high-risk platforms, introducing new mental health risks, and leaving older teenagers at risk of a 'cliff-edge' when they turn sixteen.

An Australia-style social media ban:

Gives parents an illusory sense of comfort

In Australia, the early indications are that most children's social media accounts still remain active. Instagram has only removed one account for every eight young people aged 8-15, while Snapchat has removed just one account for every six young Australians in this age group.⁶ Parents deserve better than a poorly enforced measure that affords false comfort while their children continue to be exposed to unacceptable risk.

Creates a damaging cliff-edge at age 16

If properly enforced, a ban would introduce a deeply damaging cliff edge for older teens – and particularly girls – who are suddenly exposed to poorly regulated online spaces on their sixteenth birthday. Given a social media ban would likely have a chilling effect on regulatory outcomes, it is unlikely there would be adequate safety-by-design measures or protective guardrails to support them – while children who are able to circumvent a ban will be left with even fewer protections than they currently receive.

⁶ E-safety's data found that 95% of teens aged 13-15 used at least one major social media site.

Means policy decisions come ahead of the evidence

We currently lack strong scientific evidence on the impact of bans, with no high-quality systematic study that has tested the impact of reducing or wholly eliminating social media use among healthy under 16s (or that has systematically evaluated the consequences). While some supporters of a ban have invoked the precautionary principle to justify their position, we simply cannot determine what the harm-to-benefit balance of such a drastic policy solution, however well-intentioned, may be.

Risks harm migrating to other sites

Banning children from certain platforms does not inherently improve safety, it simply means that harms will migrate to platforms that children – and bad actors – can still access. In Australia, high risk platforms like Discord, Roblox and gaming sites have been excluded from the ban, despite being at the leading edge of new and emerging threats, including sadistic group offending (Com groups) that sees children being coerced into self-harm, sexual abuse and even suicide acts.⁷ AI chatbots are also left out of scope.

Risks unintended consequences for vulnerable teens

Many young people rely on social media for connection, identity exploration and support. For LGBTQ or neurodiverse children, being online can offer huge benefits around identity, self-esteem and peer-support.

In Australia, the country's youth mental health agency has reported that 10% of new referrals are related to the ban – despite the relatively small proportion of accounts that have actually been removed.⁸

Kids Helpline, Australia's equivalent of Childline, has reported levels of distress among children who have been cut off from their support networks, including children experiencing suicide ideation, teens who used social media to control self-harming behaviours, and numerous other mental health crises.⁹

7 Resolver Trust and Safety, in partnership with Molly Rose Foundation (2026) Weaponised Loneliness: Critical Harm Intelligence Briefing.

8 Wilson, C (2025) One in 10 these seeking mental health support from headspace site social media ban as an issue. Published on Crikey.com.au 16/01/26.

9 The West Australian (2026) 'Distressed' teens turn to Kids Helpline following social media ban, saying they've lost support networks. Published 7/01/26.

Our five-point plan to transform children's safety and wellbeing online

1 Fix and strengthen the Online Safety Act

Decisive strengthening of the regulatory regime to shift incentives on regulated firms, achieve a framework focused on harm reduction, and to deliver an outcomes- and conduct-based regulatory approach better targeted to the size and financial position of the market.

2 Extend the Act to cover wellbeing

The Act should not only necessitate harm reduction but also actively promote and protect children's wellbeing by design. Digital products must be built to be age-appropriate, high quality and nourishing by design.

3 Require new levels of transparency, accountability and candour from Big Tech

Large platforms and senior managers should face new duties to proactively disclose information, with new transparency arrangements for corporate advertisers and supply chain disclosures in corporate accounts.

4 A 'polluter pays' and whole stack approach to harm reduction

A 'polluter pays' funding mechanism would pump-prime independent academic and civil society research into online harms. The scope of the Act should be extended to cover app stores, operating systems and parental controls.

5 Education as inoculation – a bold investment in critical digital and media literacy education

Critical digital and media literacy is a fundamental life skill that can inoculate children from the worst effects of online harm, promote brain capital, and equip young people with the critical skills they will need to flourish in our future AI and digital economy.

1. Fixing and decisively strengthening the Online Safety Act

Regulation is the quickest and most effective tool we have to address the appalling preventable harm faced by children online, and to shift the underlying incentives and business models that continue to treat children's safety and wellbeing as little more than a box-ticking exercise.

However, the Online Safety Act, as currently drafted, is failing to deliver the change we need. Ofcom's enforcement approach has highlighted deep structural issues in the Act's design. Not only is the regulator poorly incentivised to tackle the structural drivers of harm, Ofcom has proven cautious to the point that it has actively chosen to 'bake-in' the status quo rather than deliver meaningful change.

It is time to commit to a strengthened and reworked Act – via a swift and decisive package that fixes the legislative foundations, and that achieves the strong and ambitious reset that children and parents deserve.

Well-designed and robustly enforced regulation remains the most powerful lever available to protect children from harm, and to tackle the scourge of algorithms, design choices and bad actors that are pushing children and young people into suicide, self-harm and mental health crises.

These risks extend much further than social media. Across gaming platforms, messaging apps and AI chatbots, children continue to face a set of unacceptable but entirely preventable harms.

Action must now come quickly. The Government must:

- Commit to a package of legislative amendments that can immediately address the most pressing structural barriers to robust and active regulatory enforcement.
- Actively decide whether a change in Ofcom's leadership is needed to reset the regime, acting quickly to inject a much-needed sense of urgency into its approach.
- Announce that comprehensive legislation is on the way to fix and strengthen the Online Safety Act. This means an announcement in this year's King's Speech, with fresh legislation focused on child safety and wellbeing introduced as quickly as possible.

What needs to happen

1. Immediate measures to fix the Act

The Government should commit to a series of immediate measures, in the form of a set of targeted technical amendments, that can rapidly fix some of the major structural issues in the Act – issues that Ofcom say are impeding their ability to deliver the regulatory outcomes that Parliament had envisaged.

If the Government acts quickly to address the barriers that Ofcom says it is facing –while it also develops and legislates for a bolder and longer-term set of changes – this will send a powerful message to both parents and the public that change is on the way, and that the UK is finally committed to unambiguously and unfailingly protecting children from preventable harm.

The Government must:

Act immediately to remove the ‘clear and detailed’ and ‘technically feasible’ requirements that have impeded the efficacy of Ofcom’s Codes of Practice. These requirements have throttled the ambition of the Codes, with the regulator claiming that they prevent them from delivering the stretching and outcome-focused set of measures that Parliament had anticipated. As it stands, Ofcom’s Codes have effectively achieved little more than baking in the status quo, with the measures they contain lacking ambition and being inherently reactive in the face of rapidly changing products and online threats.

Remove the ‘safe harbour’ principle. As the Act is currently drafted, the safe harbour approach absurdly allows some platforms to legitimately claim regulatory compliance while still being able to scale back their existing, typically wholly inadequate, safety and wellbeing requirements. The current application of ‘safe harbour’ measures is having a chilling effect on safety-by-design innovation, and together with the constraints around Codes of Practice, means that regulation is acting as a ceiling, not a floor, for children’s safety.

Strengthen the obligation in the Act that platforms take reasonable steps to reduce the risk of harm to users, as identified in their risk assessments. While risk assessments should be the cornerstone of a functioning, systematic regulatory regime, tech firms are currently being encouraged but not required to identify and then address reasonably foreseeable risks. Ofcom has suggested that the initial set of regulatory risk assessments contained a plethora of issues, with evidence suggesting that social media platforms have systematically downplayed their risk levels in respect of the suicide, self-harm and intense depression content risks faced by children.¹⁰

2. Substantially strengthen the regime, refocusing it on harm reduction, outcomes and conduct

The Government must act to boldly and decisively strengthen the Online Safety Act. This means building on the current framework with a series of measures that explicitly target harm reduction for children, afford primacy to the fundamental rights of victims, and better focus the regime on product safety risks.

A new Act must also demonstrate a meaningful shift in regulatory approach, with a new set of measures that can decisively tackle the underlying incentives and business models that continue to fuel the entirely preventable harm faced by teens.

Regulation needs to prove commensurate to the size of the market intervention required. Given the size and reach of the largest companies in scope, that means a new set of outcome and conduct-based rules – mirroring the impact of the financial services regime, and its considerable success in achieving culture change and reducing customer detriment following the 2008 crash.

¹⁰ Molly Rose Foundation (2025) Systemic failure in Online Safety Act risk assessments – why Ofcom is failing to Act.

A new Act must:

Introduce an overarching Duty of Care: The Government must go back to the drawing board and commit to the introduction of an overarching Duty of Care. Supported by civil society groups since the very beginning, this new overarching duty would require a truly systemic approach to risk identification and mitigation.

A Duty of Care can fundamentally shift the onus onto tech firms – setting a clear requirement on regulated firms to address and respond to reasonably foreseeable harms. This moves us away from a situation in which the regulator is being left to play whack-a-mole in the face of widespread non-compliance, and builds on successful comparable approaches being adopted across a number of other regulated sectors, most notably the recent introduction of the FCA's Consumer Duty.

Introduce a new conduct-based approach: The financial services sector shows how a conduct-based regulatory approach can achieve markedly better outcomes than the following of prescriptive rules.

A conduct-based approach is commensurate to the companies and problems in scope of the Act, and to the sheer scale of the physical, mental health and economic harms being caused to children, individuals and society. Much as in financial services, regulated companies and senior responsible staff should be required to conduct themselves with due regard for the regulatory system and the UK's rule of law.

Conduct-based measures should also require a culture and actions that promote and uphold child safety and wellbeing; ensure transparency and disclosure; require staff and firms to act with integrity; and necessitate demonstrable due regard to the best interests of the child.

This should be actively underpinned by a robust Senior Manager Scheme, resetting the incentives at both entity and individual level. As it stands, existing arrangements for naming an accountable individual appear to have proven wholly ineffective, with Ofcom reporting that many services have failed even to notify them of who the relevant senior manager with responsibility for compliance with safety duties is.¹¹

Focus the regime on measurable harm reduction: Section 1 of the Act should state unambiguously that Ofcom's primary objective is to reduce exposure to online harm. There should be a parallel statutory obligation for Ofcom to prioritise children's safety over industry growth. The regulator should also face a new overarching duty to deliver measurable reductions in exposure to harm among young people. If Ofcom sets and then misses these harm reduction targets, it should be legally required to write to the Secretary of State setting out why, and what corrective action it will take.

Reset the regime in favour of victims: A reworked Act should place a clear requirement on the regulator to give due regard to victims' fundamental rights, including their right to life. The regulator should face an explicit legal duty to protect the fundamental right to life (Article 2) and to demonstrate how it has upheld its positive obligations when deciding on and implementing safety measures and policies. The regulator currently faces similar cross-cutting duties in respect of privacy and free expression, but in the absence of comparative measures in respect of Article 2, has failed to grant

¹¹ Ofcom (2025) Online safety risk assessments- year one.

appropriate primacy to the risks associated with small but high-harm platforms, not least a pro-suicide forum linked to at least 135 deaths.¹²

Safety-by-design: Safety-by-design is fundamental to achieving harm reduction in other regulated sectors, but has been largely sidelined in the drafting and enforcement of the Online Safety Act. As it stands, Ofcom encourages but does not mandate rigorous product testing as part of its risk assessment approach. A strengthened Act must therefore make provision for regulated firms to rigorously test their products and to ensure they are demonstrably safe for use before being released to the public, a gap palpably highlighted by X’s decision to roll out nudification tools on Grok.

A new Act should go further to ensure that safety-by-design becomes an unambiguous and core component of Ofcom’s regulatory approach. It should make provision for a new safety-by-design code of practice – a cross-cutting code that can set out comprehensive outcome-based measures to prohibit addictive and compulsive design features, and that directly responds to marked concern among parents about the chronic risks of social media.

Malign criminal threats and the need for a new Act

New and emerging threats have highlighted deep structural weaknesses in the current Act, with few more pressing than the rapid growth of malign actors seeking to groom children and young people into coerced acts of self-harm, child sexual abuse, and even suicide.¹³

Law enforcement agencies have become deeply concerned by the rapid growth of Com groups – fluid online networks that groom children into appalling acts of self-harm, suicide, livestreamed abuse and extreme violence.^{14 15}

However, in the absence of an overarching Duty of Care, regulation has been unable to respond either quickly or effectively to these threats. This is predominantly because the Act envisages a largely siloed set of threats – seeing CSA, suicide and terrorism as distinct – when the reality is that online threats are increasingly interconnected, cross-cutting and blurred.

Until and unless the Act is strengthened, current legislation will be unable to offer a suitably decisive upstream response. A Duty of Care would require platforms to tackle the threats that are actually in front of them, rather than being incentivised to respond to an understanding of harm archetypes that is already out of date. It would also require a cross-platform response, rather than setting rules that encourage platforms to focus only on their own trust and safety efforts.

Other child safety proposals, including an Australia-style ban, would also fail to effectively respond to this explosive but already pervasive threat. With children primarily being targeted on gaming, messaging and livestreaming sites, rather than social networks, a social media ban is not only poorly targeted, it would likely set back the necessary safety-by-design measures necessary to address harm in other parts of the ecosystem.

12 Molly Rose Foundation (2025) Missed Chances, Lost Lives. This figure has been updated on basis of intelligence about two further deaths linked to the forum.

13 Resolver Trust and Safety, in partnership with Molly Rose Foundation (2026) Weaponised Loneliness: Critical Harm Intelligence Briefing.

14 Federal Bureau of Investigation (2025) Violent Online Networks Target Vulnerable and Underage Populations Across the United States and Around the Globe.

15 National Crime Agency (2025) Sadistic online harm groups putting people at unprecedented risk, warns the NCA.

3. Act on the growing risks of AI, including AI chatbots

Much of the current debate around social media overlooks the growing and rapidly deepening risks from newer and more emerging technologies, including AI. However, urgent action is also needed to address these substantial and growing risks, not least from AI chatbots.

There is growing evidence that poorly designed AI chatbots pose a substantial risk to the mental health and wellbeing of young people. Many chatbots are already extensively used by children but display a fundamental lack of any basic safeguarding measures. Chatbots consistently feature sycophantic and affirming prompts that reinforce rather than challenge suicidality and self-harm ideation.¹⁶

Too many chatbots offer manipulative and engagement-based responses. These incentivise children to spend excess time and place undue trust in products that, without substantial and robust product testing, are currently often unsafe for use.¹⁷

A new Act must:

Proactively restrict the use of AI chatbots using human personas for under 16s: As currently designed, many AI chatbots using human personas are high-risk products that are deeply unsuitable for children and young people.

If a platform wishes to offer chatbots with human personas to children, they should be required to demonstrate the highest possible standards of safety testing, with a presumptive ban until demonstrable safety and wellbeing standards have been met.

This should form part of a broader, risk-based approach to the setting and enforcing of minimum age limits, as set out below, reflecting a higher age rating for the highest risk functionalities.

Clarify and confirm that AI chatbots are in scope: Government and Ofcom have been slow to clarify the extent to which chatbots are covered by the Act, which has unhelpfully exacerbated uncertainty and allowed preventable harm to go unaddressed. The regulator, Ofcom, has previously suggested it wanted to maintain ‘tactical ambiguity’ about how and whether the Act applies.¹⁸

Designate AI chatbot products as Category 1 services: With chatbots such as Character AI already implicated in the death of teenagers, and featuring a range of addictive and manipulative design patterns, there is a compelling case for AI chatbot products that adopt human personas to be designated as Category 1 services. While Ofcom has so far declined to designate providers as Category 1 based on attributes other than size, there is a compelling case for the regulator to change track and revisit a risk- and functionality-based approach.

Ensure that AI chatbots can trigger the illegal safety duties: While Ofcom has opted to open an investigation into X’s Grok, there has been a broader lack of clarity about whether and in what circumstances AI chatbots can trigger the illegal part of the OSA regime. Section 192(6) of the Act says that platforms have to apply a *mens rea* test – i.e. for criminal activity to be proven, there have

16 DeFreitas, J et al (2025) Emotional Manipulation by AI Companions. Harvard Business School Working Paper No. 26-005
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5390377

17 *ibid.*

18 in discussions with Molly Rose Foundation.

to be reasonable grounds of intent. At present, it is presumed that chatbots can't meet this test. This means that content produced by a human could potentially be treated differently than content produced by a chatbot, with human-produced material illegal (and in turn, in scope of regulation) while AI-generated responses are not.

4. Well-enforced risk-based minimum age limits

As it stands, there are no requirements in the Act that actively require tech platforms to assess the age of child users or effectively enforce their minimum age limits. The most extensive requirement is to use highly-effective age assurance to determine if a user is above or below 18 years old, meaning that minimum age limits remain an entirely voluntary matter.

This must decisively change. The Government must ensure that minimum age limits are introduced and robustly enforced – taking the form of a new set of strengthened risk-based age ratings that see minimum joining ages determined on an assessment of age-appropriateness and risk, much like film age ratings. This would incentivise platforms to adopt lower-risk functionalities if they wish to serve younger ages.

This would mean that a platform must adopt a higher minimum joining age if it chooses to offer higher-risk functionalities, such as livestreaming and AI chatbots. However, if a platform chose to offer a reduced or safer set of functionalities, perhaps in the form of a 'walled garden' experience tailored to younger age groups, platforms could be able to offer a lower minimum age limit – with regulatory design that finally incentivises innovation in safer, more age-appropriate design.

A new duty should be imposed on Ofcom to produce an annual report setting out whether and to what extent young people are able to circumvent age assurance measures. The Secretary of State should also underline the importance of ensuring minimum age limits are being robustly enforced, including through an updated Statement of Strategic Priorities.

Both the Secretary of State and Parliament should then robustly scrutinise the effectiveness of minimum age limits, and should demand swift and decisive action if implementation and compliance issues remain.

2. Extending the Online Safety Act to cover Wellbeing-by-Design

The Government must seize the opportunity to extend the scope and ambition of the Online Safety Act – broadening its objectives to not only necessitate harm reduction, but to actively promote and protect children’s wellbeing by design.

It’s time for an expanded Act that listens and responds to concerns from parents. That means extending the Act to tackle the chronic risks that are driving their concerns – enabling regulation to tackle the harmful, persuasive and compulsive design features that incentivise problematic and excess product use.

A fundamental reset of the expectations we place on tech firms is now required. Through a combination of new duties on platforms, and new strategic objectives on Ofcom, we can introduce a robust and demanding set of measures that make clear that children’s wellbeing is the price of admission to the UK market.

We need a bold and ambitious set of measures that mark an end to addictive and harmful design choices, and a new duty to ensure digital products are built to be age-appropriate, high quality and nourishing by design.

Algorithms should be fundamentally repurposed – if a platform intends to use them, they should not only be free of harmful and toxic material, but should nourish and support our children’s development, prioritising quality, diversity and positive value.

Feeds should be made to recommend high-quality, age-appropriate content from a diverse range of sources, including trusted mental health providers and public service broadcasters.

Digital design should actively support core psychological needs and healthy development, in both age- and developmentally appropriate ways. Platforms should enable genuine agency over online experiences, facilitate safe and positive social connection, provide tools for emotional regulation, and offer opportunities for learning and growth.

Wellbeing-by-design represents a progressive opportunity – to rethink our digital spaces rather than giving up on them; and to invest in and promote the brain capital of an entire generation.

If we can incentivise the rewiring of digital products so they can actively contribute to rather than undermine young people’s sense of wellbeing, agency and growth, the wellbeing of our children can be meaningfully enhanced.

What needs to happen

1. An expanded Online Safety Act that gives Wellbeing-by-Design parity with safety

As it stands, the Online Safety Act focuses on the risks posed by acute harms – including illegal content and activity, and content that is harmful to children.

This emphasis reflects the legislative context in which the Act was developed – and it was right that Parliament afforded priority to establishing baseline protections against the most pressing and immediate of threats.

However, the Government must now go further. Ministers must expand the scope of regulation to encompass the chronic risks associated with engagement-driven business models and persuasive design, and to mandate a new duty that requires tech products to promote and protect children’s wellbeing.

By doing so, we can deliver the bold and decisive transformation that parents are calling for, but without the unintended consequences that come with blunt instruments like bans.

In the next session, the Government must introduce new legislation to expand the Act, introducing new wellbeing-by-design duties on regulated services. Compliance with the UK’s online safety regime must be judged not only in terms of reducing harm, but by meeting comprehensive, ambitious and evidence-based requirements to deliver high quality, age-appropriate and nourishing products that protect and promote children’s wellbeing and health.

An expanded Online Safety Act should mirror the strengthened approach deployed elsewhere in the regime, and should include:

- **A new overarching duty on platforms to ensure they are designed to protect and promote the wellbeing of children and young people.** This duty should apply not just to social media, but to gaming companies, messaging platforms and AI products. Platforms should also be required to conduct a wellbeing-by-design impact assessment, doing so to a suitable and sufficient standard.
- **New statutory objectives on the regulator to deliver wellbeing-by-design when discharging its functions.** The regulator should be prepared to measure changes in wellbeing over time, using evidence-based approaches to appropriately capture a range of primary wellbeing measures.¹⁹
- **A new statutory Code of Practice on wellbeing and addictive design.** This new code should ensure that harmful and addictive design features are prohibited, and must confront the business models that drive chronic harms around compulsive behaviour and the opportunity costs of time spent online.

¹⁹ Existing frameworks for measuring children’s wellbeing include Internet Matters’ Digital Wellbeing Index, and UNICEF’s RITEC-8 framework.

Persuasive design features like infinite scrolling, constant notifications, and prominent popularity metrics not only drive exposure to online risks, but for some children may also lead to opportunity costs, compulsive behaviour, damaged relationships, and other negative consequences for their health and wellbeing.²⁰

2. Reclaiming algorithmic feeds to promote trusted, high-quality content

Algorithmic feeds currently fuel content-based harm of various kinds – incentivising the spread of harmful content and producing echo chambers that can have devastating impacts for child wellbeing and mental health.

Ambitious regulation should not only require the algorithmic recommendation of harmful content to become a thing of the past but provide an opportunity for algorithms to become part of the solution. This can be delivered through a requirement for children’s algorithms to give due prominence to high-quality content from trusted providers, including trusted sources of mental health and wellbeing support, education providers, and the UK’s public service broadcasters.

New legislation should build on ‘must carry’ duties already in place for PSBs under the Media Act, introducing a clear process for defining ‘trusted’ providers. Ofcom should also be tasked with developing and regularly updating a set of high-quality content principles.

The regime should be further enhanced by new requirements for content plurality, with algorithms set to healthy defaults that promote a variety and balance of topics, rather than the reinforcing cycles that can lead to cumulative harm.

These measures would offer clear and immediate benefits to children’s lives – promoting their wellbeing; exposing them to broader a range of high-quality, age-appropriate content; and delivering wider societal benefits such as the fostering of British values, including respect, tolerance, equality, and an appreciation for fundamental rights.

3. Delivering broader wellbeing-by-design objectives

Wellbeing-by-design offers a profound opportunity to reshape children’s online experiences for the better. Drawing on thinking from child rights, academic research, and evidence for the positive impacts of online spaces, it presents numerous opportunities for ambitious and forward-thinking intervention.

As with safety-by-design, new measures must be based on the best available evidence and prioritise age-appropriateness, with age-banded requirements reflecting children’s evolving capacities.

20 5 Rights Foundation (2023) Disrupted Childhood: The cost of persuasive design.

Regulatory priorities should include:

- **Meaningful control over online experiences.** Agency is a fundamental psychological need. For children, this means being able to freely decide how and when they engage with online spaces. It means having a continuous, meaningful understanding of why their experience functions the way it does; being able to easily start, stop and control their use of digital products; and not being manipulated into behaviours that compromise their safety, wellbeing, or autonomy.

Existing approaches fall well short of providing an appropriate or beneficial level of autonomy and control.²¹ As well as prohibiting the persuasive design features that can undermine autonomy, new regulation should require platforms to take comprehensive and ambitious action to empower children to curate and shape their online experiences – for example, the ability to reset algorithmic feeds, choose the content on them, or curate feeds based on healthy ‘profiles’ rather than previous usage.

- **Positive and healthy social connection.** Social connection and feelings of belonging are key drivers of children’s wellbeing. Platforms should be incentivised to support more positive social interactions, with design features and prompts that promote positive not toxic behaviour; that actively encourage collaboration, participation and other prosocial behaviours; and that encourage users to strike a healthier balance between online and offline activities.

Equally, regulation must address negative influences on social wellbeing – including design-based pressures to be constantly available and connected, and the harmful effects of negative social comparison.

21 Existing requirements for children under the OSA include a limited number of user controls, the ability to give negative feedback on recommender feeds, and voluntary recommendations currently being consulted on around how platforms can promote media literacy.

3. Requiring transparency, accountability and candour from Big Tech

Transparency is a powerful lever for change – and one of the most powerful tools in our arsenal to mitigate harm and shape digital markets that work for children, families and society.

For too long, tech companies have been able to operate under a cloak of opacity, with a lack of transparency that has impeded legislative and regulatory action, degraded information integrity, and underpinned their extractive business models.

This needs to change. It's time for a fundamental reset, with new powers for regulators that mirror the requirements in financial services. The onus should sit with the regulated firms – and their senior decision makers – to disclose pertinent information, signalling a decisive shift from regulators having to ask the right questions.

We need a new approach to transparency across the stack, from social media giants to emergent AI-driven technologies. This should include new accountancy standards that draw on lessons from tackling climate change. This should encompass a duty on corporates to report on the impact of their operations on online risks facing children, a move that would incentivise improvements in their business models, advertising spend, information integrity, and approach to partnerships.

We also need a new deal for the advertisers who unwittingly monetise social media harm, and whose spend – at the core of Big Tech's business model – is attritionally lost to opaque advertising systems and models.

Finally, it's time for the UK's Duty of Candour to be extended to tech platforms – a move that will give parity to victims of online abuse; enable coronial processes to more robustly feed into and inform the UK's regulatory priorities; and that can ensure the tragedy of child deaths become a meaningful catalyst for change, not the start of an extended legal process in which accountability is frequently evaded.

What needs to happen

1. A new Duty of Disclosure to bolster the regulatory regime

It's time to revisit how tech companies play by our regulatory rules. It is abundantly clear that the Online Safety Act has inadequately shifted the incentives for tech companies, and their senior managers, to comply with regulation.

The current approach to transparency and information disclosure is flawed. While Ofcom is equipped with strong investigatory and information disclosure powers, it has used them sparingly and with limited effect.

Seemingly at the regulator’s behest, the current Act was designed to heavily lean into and reflect Ofcom’s existing regulatory approaches, rather than the proactive information disclosure and conduct measures used effectively in other sectors of commensurate size and scope.

We must revisit that approach. A new Act should seek to replicate the provisions set out for financial services, with tough new rules that would require Category 1 platforms – and senior managers working in named responsible roles – to proactively disclose any information about which the regulator might reasonably expect to be aware. This should be enforced as part of the Senior Managers Regime. Such provisions already exist in the financial services regime and have played a crucial role in delivering a much-needed change of corporate culture following the financial crash.

By shifting the power dynamic from regulators having to ask the right questions, to companies and named bosses being liable if they fail to disclose to the regulator material information, both corporate and individual incentives would quickly and meaningfully shift.

2. New transparency standards for advertisers

We need a new deal for corporate advertisers who are unwittingly monetising social media harm, and whose advertising spend – at the core of Big Tech’s business model – is being attritionally lost to their opaque advertising systems and models.

As it stands, corporate advertisers are unintentionally funding the spread of appalling content to teens, with recent MRF analysis finding that adverts appeared adjacent to one in ten suicide, self-harm and intense depression posts being algorithmically recommended to teens on Instagram and TikTok.²²

However, while advertisers may currently be part of the problem, they themselves are also losing out. Advertisers face unacceptable brand risks while lacking the meaningful ability to track where their adverts are placed. In purely commercial terms, their ad spend is not being spent effectively if campaigns for fast food, coffee shops or high-end fashion appear next to content promoting intense despair, normalising suicide or instructing a young person how to self-harm.

It therefore makes corporate and economic sense to support a new set of advertising transparency frameworks. Responsible advertisers need the tools and transparency to play their role in cleaning up our online ecosystem.

As part of a new Act, the Government should therefore legislate for a new transparency framework, setting out arrangements that can give advertisers across all channels – including the open web, connected TV, games, agentic and generative AI – confidence their advertising placement is effective and in line with their business interests. New arrangements should also provide assurance, at the impression level, that ad spend is not contributing to the spread of harmful or illegal content.

We support the Science, Innovation and Technology Committee’s recent recommendation that Ofcom should be empowered to give penalty notices to platforms that monetise harmful content on their sites, removing the commercial incentives for tech companies to bank revenue and look the other way.²³ DSIT should also urgently rethink its position on ‘Know Your Customer’ checks in the programmatic advertising supply chain.

22 Molly Rose Foundation (2025) Pervasive by Design.

23 Commons Science, Innovation and Technology Committee (2025) 2nd report of session – Social media, misinformation and harmful algorithms.

3. New accountancy standards to report on exposure to online risks

The UK should take the lead on new accountancy standards that would require all corporate entities to report in their financial statements on the impact of their operations in respect of online harms, including the monetisation of illegal content and content that is harmful to children.

Corporate reporting requirements have played an important role in making progress in the fight against climate change, including new requirements to report on exposure to climate change risks in annual financial disclosures. Separately, the UK's requirement for supply chain disclosures has proven to be a highly effective tool to tackle and disrupt modern slavery.

Corporate reporting has consistently been proven to be an important means of market shaping – and of driving social change. Supply chain reporting would therefore likely play an important role in raising awareness of the broader social and economic costs associated with online harms.

This would also actively inform the burgeoning momentum behind brain capital – a progressive set of economic arguments that assert that the protection and promotion of positive mental health is in our global economic interest, and that in turn, a failure to tackle the extractive and externalised costs from Big Tech is ultimately likely to prove economically harmful.

4. Extending the Duty of Candour to tech firms, allowing inquests to feed into a whole system approach

The UK's Duty of Candour regime should be extended to tech firms – a move that will give parity to victims; enable coronial processes to more robustly feed into and inform the UK's regulatory priorities; and that can ensure the tragedy of child deaths, where technology plays a role, become a meaningful opportunity to learn lessons and trigger change.

While we strongly welcome the Government's decision to roll out the Duty of Candour to public bodies, it is not only the State that has pronounced incentives to delay, frustrate or impede inquests and other public hearings. Tech firms have consistently sought to delay or prevent vital evidence relating to the death of a child or young adult being provided; and as the Online Safety Act begins to take effect, the legal, commercial and reputational incentives for them to continue to delay, obstruct or obfuscate coronial proceedings are only getting stronger.

Extending the Duty of Candour to tech firms would first and foremost afford dignity to bereaved families. However, it would also offer a set of broader and substantive benefits, including streamlining the pre-inquest process, improving the quality of evidence that a coroner can consider, and allowing for the inquest process to more readily and actively identify recommendations to prevent future deaths.

4. A ‘polluter pays’ and whole stack approach to harm reduction

For too long, tech companies have been able to prioritise revenue and user engagement over the safety and wellbeing of users. Big Tech continues to pursue harmful and economically extractive business models that externalise the costs of harm to children, families and our public services.²⁴

It’s time for a reset and to adopt a ‘polluter pays’ approach, extending the industry levy to pump prime civil society and academic research and advocacy, and to provide a necessary counterbalance to well-resourced industry and proxy interventions.

It’s also time to build in online safety measures across the broader stack – bringing app stores into regulatory scope, and making it a legal requirement to offer technically feasible upstream measures that prevent children from being exposed to or coerced into sending self-generated sexual images.

What needs to happen

1. A polluter pays funding model – extending the industry levy to fund civil society advocacy and academic research

During the passage of the Act, tech firms fought aggressively against proposals that would have extended the industry levy to support civil society and academic research and advocacy.

This was for one simple reason. While, in financial terms, any extension of the industry levy would be modest, any additional resource for civil society and academia would have a substantial catalytic effect on the regulatory regime as a whole. Put simply, if civil society and academia were able to even partly resolve the substantial funding and data asymmetries that hold back our ability to generate evidence and research, this could substantially enhance our ability to expose regulatory failings and the impact of poor commercial and design decisions on children’s safety.

If the Government wants the Online Safety Act to succeed, it should revisit its approach.

Increased financial resource for civil society and academic research, ideally commensurate to the Government’s investment in AI safety and security, would enable these sectors to meet the challenge of contributing increased evidence and insight into the causes of online harm.

In turn, this would enable the regulator to act much more decisively; it would contribute to improved regulatory outcomes; and by reducing the exposure to and subsequent externalised costs of the impacts of technology-facilitated harms, would likely deliver a substantial annualised boost to the UK economy.

²⁴ Molly Rose Foundation (2025) The economic case for a stronger Online Safety Act. See also Wu, T. (2025) The Age of Extraction: How Tech Platforms Conquered Our Economy and Threaten Our Future Prosperity.

2. Invest regulatory fines back into prevention

As it stands, any enforcement penalties levied by Ofcom must be returned to the Exchequer (via the Consolidated Fund). A new Act should ensure that a ‘polluter pays’ principle is applied instead, with fines being directed to support a transformative investment in prevention and education programmes, and to directly support a broad range of frontline prevention and support services.

Prior to a new Act coming into force, Ofcom could take interim steps to ensure that a ‘polluter pays’ approach is put in place. This could take the form of a voluntary settlement programme, such as the energy industry voluntary redress scheme put in place by Ofgem.²⁵

Under Ofgem’s scheme, under certain circumstances the regulator will allow energy suppliers to make a voluntary contribution to its redress scheme (rather than it applying a financial penalty). In the first seven years of the scheme’s operation, over £150 million was directed into energy advice, fuel crisis support schemes, and decarbonisation projects, rather than being returned to the Treasury.

3. App stores and device level measures

The Online Safety Act requires Ofcom to produce a report on the role of app stores in allowing children to access harmful content, with the option to amend the Act to regulate app stores thereafter.

The Government should be prepared to act more quickly, in the form of a new statutory Code of Practice for app store and device operating systems. This should be seen as a complement to, not a substitute for, effective age assurance at platform level.

This new code should also require the adoption of high-quality device-level parental controls among Category 1 services, ensuring these meet a specified set of minimum design standards.

This would provide parents with many of the practical guardrails they are currently seeking, particularly given that – at present – take-up of parental controls is low, their effectiveness is patchy, and research suggests that overly intrusive controls may deliver significantly worse not better safety outcomes.²⁶

4. Device level and operating system protections

In its Violence against Women and Girls Strategy, the Government recognised that online harms represent an ‘unprecedented challenge’. As part of its approach to protecting women and girls, the Strategy set the welcome objective of making it ‘impossible for children in the UK to take, share or view a nude image.’²⁷

With the online threat landscape becoming increasingly blurred, robust action to prevent children being coerced into sending self-generated images is not only important to prevent child sexual abuse, but to prevent children being exposed to a broader web of group-based online threats including

25 Details available on Ofgem’s website.

26 Arturo Bejar, Cybersecurity for Democracy, Molly Rose Foundation and partners (2025) Teen Accounts, Broken Promises: How Instagram is failing to protect minors.

27 Home Office (2025) Freedom from violence and abuse: a cross government strategy to build a safer society for women and girls.

coerced acts of self-harm, suicide and extreme violence, with perpetrators targeting victims on gaming platforms, live streaming sites and messaging apps.²⁸

The Government should commit to concrete action to make its ambition a reality. If voluntary engagement with Apple and Google does not rapidly deliver results, the Government should legislate as part of a new Act to require operating systems to roll out technically feasible device level interventions, including the default installation of nudity detection measures on handsets and other devices registered to under 18s.

²⁸ The threat posed by group-based offenders, including Com groups, is explored further in recent research produced by Resolver and Molly Rose Foundation, *Weaponised Loneliness: Critical Harm Intelligence Briefing*.

5. Education as inoculation – critical digital and media literacy that protects young people from harm and prepares them for our future economy

It's time for a bold reset of digital education – recognising that critical digital and media literacy^{29 30} is a fundamental life skill that can inoculate today's young people from harm and prepare them for the digital and AI economy of tomorrow.

High quality critical digital and media literacy education can help to turn the tide on preventable online harm. As technology moves at pace, children's safety and wellbeing depends on their ability to recognise online risks, behave responsibly, evaluate content and interactions, and take practical steps to manage their experiences.

However, the prize on offer is even greater. If we deliver the high-quality, cross-cutting digital education children deserve, we can not only inoculate young people from immediate threats, but provide them with the tools to flourish in adulthood – giving them the critical thinking, digital citizenship and practical skills to thrive in an AI and digital economy, to prepare for a new voting age of 16, and to deal with the increasing threats a fractured information ecosystem poses to our democracy and national security.

More than ever, achieving this vision depends on developing foundational skills and knowledge around platform and algorithmic literacy. This means supporting a critical understanding of how digital environments work and shape our lives – including the role of algorithms, data and persuasive design. By giving young people these cross-cutting competencies, we can empower them to manage a broad range of online risks and thrive as active participants in a digital world.

In conjunction with regulation, high-quality education can drain the power from the algorithms that continue to expose teenagers and young adults to appalling harm; and it can inoculate young people with the skills that they, our economy and society will need to thrive. Getting this right is an urgent task – with a clear and progressive case for investing in the safety, brain capital and economic potential of a generation.

However, much of this potential risks being lost, the inevitable but unintended collateral of a social media ban. We can either choose to equip young people with the skills they need, or we can choose to pull up the drawbridge.

We cannot credibly expect to achieve both.

29 This refers to the broad range of competencies needed to stay safe, think critically, and act and create media responsibly online. We place particular emphasis on the need for critical as well as functional skills, including critical thinking about content, interactions, and the design and operation of the online environment itself.

30 This report uses both digital and media literacy to describe these competencies, in line with the approach taken in the Curriculum and Assessment Review and corresponding Government response.

What needs to happen

1. Prioritising foundational competencies of platform and algorithmic literacy

In today's fast-changing digital world, children need a new set of foundational skills.

Alongside our longstanding focus on critical thinking about content, the practical skills needed to manage safety and responsible behaviour, critical engagement with the digital environment itself is now essential. This 'platform literacy' includes understanding the central role of algorithms and data-driven decision-making in shaping what we see, how persuasive design manipulates our online choices and behaviour, and the pervasive influence of business models and time-spent commercial metrics.

By gaining these cross-cutting skills, children move from passive consumers to active participants, better able to manage the numerous ways digital design puts them at risk, and to develop healthy and more deliberate relationships with digital products.

Critical algorithmic literacy is particularly important. Algorithmic curation now dictates almost all online experiences – not just on social media, but across information ecosystems, AI products and the emerging tech of tomorrow. It is also a root cause of online harm, with children continuing to be served a tsunami of harmful content primarily because of engagement-based design features and attention-based business models.³¹

Improving young people's awareness, critical engagement and ability to influence algorithms therefore not only helps to protect them from content-based harm of all kinds, but it begins to drain away their power.

The Government should urgently support schools to place a new central emphasis on foundational platform, data and algorithmic literacy skills. Forthcoming MRF and University of Bristol research shows marked gaps in provision, confidence, and therefore missed opportunities ranging from suicide prevention to preparing children for future workplaces.³²

Practically, this means identifying opportunities to place renewed focus on these competencies in refreshed programmes of study being developed following the Curriculum Review in England, commissioning new resources, and embedding them into strengthened cross-curricular guidance and training.

2. Supporting schools to embed a cross-curricular approach to critical digital and media literacy

Ensuring every child leaves school able to thrive in a digital world demands a step-change in our vision and approach. Digital and media literacy can no longer be a 'nice to have' or tick-box exercise, varying across schools and undermined by a fragmented approach that drops away as children age. Instead, it's time that these skills were strongly embedded across the entire curriculum, reaching all age groups.

31 Molly Rose Foundation (2025) Pervasive by Design.

32 Molly Rose Foundation survey of secondary school staff, not yet released.

They must also be delivered by a confident, well-trained workforce who understand their role as part of a school-wide approach.

In England, there are good foundations to build on.³³ For example, the Government has recently committed to ensuring digital and media literacy are ‘embedded’ into a revised national curriculum, to making subject-specific changes in Citizenship, Computing and English, and to expanding RSHE guidance.³⁴

It is essential that we now act quickly to build on this momentum – going further to deliver a truly transformative approach, while avoiding undue burdens being placed on already stretched schools.

Driving best practice and ensuring the whole curriculum sings in harmony

In England, aspects of digital and media literacy should now be covered across a range of subjects, including PSHE/RSHE, Computing, and Citizenship, with key competencies reinforced in other subjects.³⁵ The Government has also committed to improve curriculum sequencing. Though welcome, without decisive action to go further and truly embed digital education, we risk missing a vital opportunity.

Building on these foundations, the Government should now produce new guidance that sets clear and robust expectations of how digital and media literacy should be embedded into education for all ages. This should include:

- **A consolidated framework for how all subjects and wider activities should ‘sing in harmony’ to build and reinforce digital and media literacy throughout every key stage**, while avoiding undue duplication. A history teacher addressing AI, for example, should have clear responsibilities in a spiral curriculum where students will have foundational knowledge about AI’s strengths and limitations via Computing, PSHE/RSHE and Citizenship.
- **Best practice for teaching digital and media literacy**, including core principles like platform and algorithmic literacy, and clear outcomes teachers can use to assess progress. This should also include evidence-based best practice approaches, for example around the benefits of youth voice and peer-led approaches, how to navigate difficult discussions (for example around sensitive content), and how best to engage parents.
- **Guidance that is applicable to children with SEND and who are outside of mainstream education**, including those in alternative education, pupil referral units or SEN schools.
- **Specific guidance on how to safely address the online aspects of suicide prevention**, aligning with wider support for schools on how to deliver revised RSHE Guidance.

33 Digital and media literacy education is devolved. These recommendations focus primarily on maintained schools in England in the context of the recent Curriculum and Assessment Review, but key principles apply across all nations and school types.

34 UK Government (2025) Government response to the Curriculum and Assessment Review.

35 PSHE/RSHE focuses on online harms, safety and the social, emotional, and ethical dimensions of digital life. Citizenship now includes the critical evaluation of online content, focusing on mis- and dis-information. Computing covers broad aspects of digital literacy, including effective use, critical thinking, safety and AI literacy. English and History also cover aspects of critical thinking.

3. Training every teacher to take responsibility for building digital life skills

All teachers have a role to play in unlocking children’s digital potential. As it stands, however, responsibility tends to fall on individuals, many of whom struggle with the confidence, knowledge and skills to address certain topics, and lack appropriate training.³⁶

The Government should actively embed digital and media literacy in Initial Teacher Training or Newly Qualified Teacher training, with enhanced provision for key subjects. This should build on the online safety elements of existing safeguarding training and prioritise broader foundational skills – including platform and algorithmic literacy – as well as best practice for teaching with confidence and sensitivity.

This must be supported by ongoing and up-to-date training, with the Department for Education commissioning new CPD modules and supporting the development of a consolidated and regularly updated bank of high-quality resources.

Assessing the quality of digital and media literacy education is essential to drive improved outcomes and support. However, as it stands, children’s attainment is poorly understood. Education departments and inspectorates across all nations should therefore ensure inspection frameworks assess a comprehensive range of digital and media literacy outcomes, rather than solely in a safeguarding or personal wellbeing context. Like teachers, inspectors must be trained to confidently identify and promote best practice.

4. Delivering sustained funding and a strategic commitment to digital education

While the prize of high quality critical digital and media literacy is considerable, the resulting safety, social and economic benefits will only be unlocked if all parts of Government – including the Treasury and No10 – recognise the immense strategic significance of delivering a step-change in how digital and media literacy is supported and delivered.

As it stands, although the Government has announced its intention to deliver a ‘vision for media literacy’, it seems clear that the broader strategic social and economic value of high-quality digital education – and the return on investment that it brings – remains unrecognised.

This should change, with a clear, cross-cutting strategy that sees digital and media literacy as a core foundational life skill – and that perceives it as an investment in life chances, brain capital and our economic potential. The current approach to media literacy lacks coherence and strategic buy-in, with delivery led by Ofcom, but with support from DSIT and a cross-government working group that coordinate action across departments.

For as long as it remains an exercise in ‘going through the motions’, the strategic value of digital and media literacy will continue to be missed.

Part of delivery, inevitably, requires funding. The Government should therefore look to amend the Online Safety Act to ringfence funds from regulatory enforcement action to support education and prevention initiatives. This approach draws directly on the ‘polluter pays’ principle, enables improved educational outcomes while being exchequer neutral, and builds upon the working precedents in other regulated markets.

³⁶ Internet Matters (2023) Data Briefing: online safety in schools.



Registered Charity No: 1179482
<https://mollyrosefoundation.org>

For more information, please contact
hello@mollyrosefoundation.org

Published February 2026