

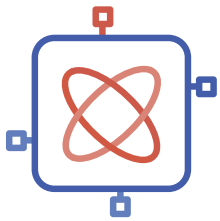


**Computer & Communications
Industry Association**

Open Markets. Open Systems. Open Networks.

 ccianet.co.uk

POLICY MEMO



UK Government Proposals on AI Chatbots

September 2026



2026 Policy
Memo

POLICY MEMO

UK Government Proposals on AI Chatbots

This paper is a discussion document prepared for officials ahead of further announcements on AI chatbots. It explores where the sets of measures announced in June and July 2026 would produce outcomes the Government is unlikely to intend, and identifies the policy decisions still open.

In summary, this paper makes two broad points. The first is about scope: the evidence that the Government appeals to concerns a narrow class of services designed to simulate ongoing relationships, while the measures reach every chatbot a child or young person touches, including general-purpose assistants, educational tools, workplace software and emerging agentic products. The second is about the policy mechanism itself: a hard timer-based break does not respond to individual contexts, so it interrupts the interactions that matter most alongside the ones the Government is worried about.

Proposed measures

The measures proposed by Government apply to everyone under 18, including 17 year olds, and none can be switched off by the young person. This is stricter than the social media regime set out in the same document, where 16 and 17 year olds receive default settings they are expressly permitted to change (CP 1643, p.11). A 17 year old can switch autoplay back on but the same 17 year old cannot switch off mandatory chatbot breaks.

- ✦ The harm the Government articulates is emotional attachment formed through prolonged conversation with services designed to simulate a relationship (CP 1643, p.19). Whilst that may be a legitimate area to consider policy interventions, it does not describe a child checking a chemistry answer on a general purpose assistant, using a wearable to read a menu aloud, or drafting a customer email on a workplace tool. Currently, the proposed measures do not distinguish between these scenarios.
- ✦ The Government response records that respondents were clear that chatbots offer children benefits “around education, learning and creativity” (CP 1643, p.18). It accepts that chatbots simulating a human persona “can have benefits in some supervised contexts”, giving foreign language practice as its example (CP 1643, p.19), and quotes the Centre for Information Policy Leadership on educational and accessibility-focused tools supporting “learning and inclusion” (CP 1643, p.19).
- ✦ The June statement is more specific on what children actually do with these tools. Parents, children and young people ranked help to find information and support for learning as the two highest benefits chatbots offer (Progress statement, p.29). General purpose chatbots are “the chatbots most used by children” (Progress statement, p.30). The statement quotes Amelia, 15, on managing her dyslexia: “If I didn’t have it, I’d probably struggle quite a lot” (Progress statement, p.30), and a 15 year old on asking a chatbot to explain a maths question the teacher had not explained properly (Progress statement, p.30). The National Youth Agency told the consultation that chatbots can

help young people with special educational needs and disabilities reach their full potential (Progress statement, p.30).

- ❖ Chatbot breaks are made mandatory for everyone under 18, with no default option, no differentiation by age, and no proportionality assessment of any kind (CP 1643, p.18-19, 21). This is a departure from the approach taken by other online safety measures. The Government does not explain why default settings are effective enough to make forced-on curfews disproportionate for a 17 year old on social media, and not effective enough for the same 17 year old on a chatbot. Nor does it explain why the autonomy at 16 and 17 that it recognises in chapter 2 does not then carry over into chapter 3.

What the measures would mean in practice

The scenarios below are illustrative rather than exhaustive. Each is drawn from a use case the Government's own documents record, or from a straightforward application of the measures as announced.

A. Mandatory Breaks

The Government response requires AI chatbots to introduce mandatory breaks for all children under 18 (CP 1643, p.18, 21). Frequency and duration are not set and are to be settled with experts including the Expert Panel chaired by Professor Russell Viner (CP 1643, p.19). The measure is justified by concern about prolonged engagement and emotional attachment (CP 1643, p.19). We believe it applies to all chatbot use, including the educational use the same document identifies as a benefit, and would welcome clarification if that is not the intention.

Two questions arise here which are separate and distinct. The first is which services should be in scope. The second is whether a hard break is the right instrument for those that are.

On scope, the Government's own evidence describes how 18% of 16 to 21 year olds responding to the children's consultation said that young people see chatbots as a source of emotional or mental health support (CP 1643, p.18). A policy intervention designed around that finding would target services built around emotional interaction but currently the measures as described so far reach many other cases as well.

On the instrument itself, a time limit takes place with no account for the context of what is being discussed. It cannot distinguish revision from reflection, and it cannot tell whether the exchange it is ending is one in which a child is being directed to a verified and good external source of support (eg. Childline or the Samaritans). Where the Government's concern is the length of an interaction, graduated tools already in use across the market address it without the risk of ending an interaction at the point it matters most.

We are not aware of another jurisdiction that mandates hard breaks on general purpose AI assistants for minors. Where governments have acted, measures have generally been directed at companion services specifically.

- ❖ A 16 year old is halfway through a step-by-step explanation of a quadratic equation on a general purpose assistant the night before a GCSE paper. The mandatory break ends the session and the working is lost. This is the use case that the Government itself

documents (Progress statement, p.30). A child with a tutor or knowledgeable parent at home has an alternative, but a child without one does not.

- ❖ A dyslexic 15 year old uses a chatbot as a reading aid for a set text. The break interrupts them repeatedly, and it never interrupts classmates reading the same book on paper. The Government's own case for chatbot benefits is a dyslexic student (Progress statement, p.30), and the response quotes evidence on accessibility-focused tools supporting inclusion (CP 1643, p.19).
- ❖ A visually impaired 16 year old uses AI-enabled smart glasses for live scene description, navigation and reading text aloud. A hard break switches off an assistive function in the street. This is neither screen time nor emotional attachment and the Royal Society of Blind Children has piloted exactly this use.
- ❖ A 14 year old discloses bullying to a chatbot after weeks of silence. The enforced pause lands mid-conversation and they do not start again. The June statement records that children value a non-judgemental and easily accessible way to discuss their feelings (Progress statement, p.31). Worse, a break that kicks in as a service is signposting a helpline removes the child from the referral rather than from the conversation, producing the opposite of the safeguarding principles that are intended.
- ❖ A 13 year old with autism rehearses a difficult social situation before school and the break cuts the rehearsal short. The Government's response accepts that persona-based chatbots have benefits in supervised contexts (CP 1643, p.19) without providing any mechanism by which a supervised context is treated differently.
- ❖ A 16 year old debugging code in an international school hackathon faces mandatory pauses that pupils abroad in the same competition do not.

B. Agentic and Background Use Cases

The measures proposed so far by the Government assume a conversational model where a child asks, the service answers, and a session can be timed. However, real world uses are moving beyond that model already. Assistants increasingly carry out tasks on a user's behalf, monitoring for an event, working through various stages of research, holding a schedule, running while the user is not at the screen. If the regulations define a chatbot by its natural language interface, these tools are in scope, and a mandatory break does not interrupt a conversation but rather ends the entire working process.

If background activity counts toward the timer, a child is penalised for use that is not truly engaged use. The distinction between interactive and functional use has to be settled before the regulations are made, not reviewed afterwards.

- ❖ A 17 year old asks an assistant to alert them when course registration opens. The break fires while the task is running and the window is missed.
- ❖ A 16 year old sets a revision timetable and a reminder for a medical appointment. Neither is delivered because the daily allowance is spent.
- ❖ A 17 year old runs background research for an Extended Project Qualification over the course of the week. Progress is lost each time the timer expires, and the nature of the work means it cannot just be resumed where it stopped.

Rules written for a 2026 conversational product will bind 2027 mainstream agentic ones. We would encourage the department to address the distinction with some urgency.

C. Age Assurance and the Threshold Gap

The June statement promised that access to general purpose AI models would not be age gated (Progress statement, p.30, restated s.5.1, p.32). The Government response requires mandatory breaks for all under 18s (CP 1643, p.21). A service cannot impose breaks on under 18s without first establishing which of its users are under 18. The response does not restate the no-age-gating promise, does not withdraw it, and does not address the tension.

More concretely, the age assurance work programme is scoped to the wrong threshold. The department asked Ofcom for a rapid assessment of what highly effective age assurance looks like “for determining whether someone is over 16”, to be published in October (CP 1643, p.23). The circumvention duties are framed in the same terms, requiring companies to take robust steps to prevent under 16s circumventing the new rules (CP 1643, p.24). Every chatbot measure, however, operates at 18. There is no work programme establishing what highly effective age assurance looks like at the 18 threshold, no timetable for one, and no acknowledgement in the response that a second threshold is required at all. The chatbot regulations are due within a year of publication (CP 1643, pp.4-5) and it appears that the assurance methodology they depend on is not being developed.

- ✦ A 15 year old uses a general assistant for homework, the use the Government promised would not be age gated. To apply her mandatory breaks the service must first establish her age. Age profiling of every user would need to occur.
- ✦ A 12 year old uses a homework helper that previously required no account. Break enforcement ends anonymous use, and a child too young for an email address must be age assessed to ask about, for example, photosynthesis.
- ✦ A sixth-former uses an open source model for a computing project. Compliant UK services carry verification requirements, but non-UK and self-hosted alternatives carry none. The most curious pupils may be pushed towards the least safe tools.
- ✦ A family shares one laptop. Age assurance keyed to the account means a parent is repeatedly challenged to prove their age, or a 15 year old inherits an adult session. The rule fails in both directions.

D. Mental Health and Wellbeing Advice

The Government response commits to stronger action on chatbots posing risks to children through “harmful, inaccurate or unverified mental health advice”, and to considering a full range of options including banning certain services where other regulatory and safety mechanisms are insufficient (CP 1643, p.18, 20). It will work with DHSC, NHS England and the MHRA (CP 1643, p.20).

The standard for this is not yet defined and the regulatory route is potentially narrower than the commitment it is meant to deliver. The MHRA can act where products within its remit pose risks to public health, which it describes as covering chatbots that claim to diagnose, treat or manage a mental health condition (CP 1643, p.20). A general purpose assistant answering a bereaved 14 year old makes no such claim and falls outside that remit entirely.

A service that claims to provide mental health care should be regulated as such accordingly, and the MHRA is the right body for that. However, a general purpose assistant that responds supportively to “I’m stressed about my exams” or similar inputs makes no clinical

claim. Undefined “unverified advice” scenarios provide problematic areas for providers looking to assess the risk and potential regulatory requirements.

It is also worth noting what already exists. Most general purpose services operate crisis detection and redirection to services such as the Samaritans and Shout, refusal policies on self-harm and disordered eating content, parental notification where a conversation suggests risk, and time and content controls within supervised accounts. These are the mechanisms the response contemplates as alternatives to a ban (CP 1643, p.20) and the Government’s starting point should take account of the measures that are already in place and the evidence of their effectiveness. June statement’s balancing commitment, that children should still be able to access and benefit from trusted support (Progress statement, p.33), does not appear in the subsequent July statement from Government.

- ❖ A 16 year old asks what to say to a friend who has told them she is self-harming. A provider configured against an undefined unverified-advice standard refuses, and does not signpost Childline or any other source of help.
- ❖ A teenager with exam anxiety asks for a breathing exercise before an oral, and is declined on the basis that this is wellbeing advice the service no longer risks giving.
- ❖ A 15 year old asks the symptoms of depression for a biology or psychology project and is refused as the provider cannot distinguish coursework from crisis.
- ❖ A 17 year old carer asks how to support a parent with bipolar disorder. Refused as mental health advice.
- ❖ A bereaved 14 year old asks how long grief usually lasts. Over-compliance treats an emotional question as a medical/health question.

E. The 18+ Gate and Definitional Lines

Under 18s are to be prevented from accessing chatbot services that primarily offer sexualised content (CP 1643, p.7). The “primary purpose test” is applied by the provider, in advance, against an undefined line.

- ❖ A 16 year old preparing for a GCSE speaking exam practises a roleplay about asking someone out, which is the kind of everyday social situation language courses use because it is exactly what a learner is encouraged to learn. The chatbot is configured against an undefined intimate functionalities boundary and declines. The Green paper identifies foreign language practice as precisely the supervised context in which persona-based chatbots have benefits (CP 1643, p.19). The primary purpose test cannot cleanly exempt these services, and the breaks requirement applies to them regardless of the exemption question.

F. Enforcement and Circumvention Duties

The Government response confirms it will not ban or age-gate VPNs (CP 1643, p.22). It places the onus on in-scope services, which must take robust steps to prevent, detect and respond to underage circumvention (CP 1643, p.22), and has commissioned Ofcom to research how services can detect VPN traffic and prevent children using VPNs to bypass safeguards (CP 1643, p.24). Ofcom’s age assurance assessment is due in October 2026, and its circumvention research and report is not due until spring 2028 (CP 1643, p.9), after the chatbot regulations are expected to be made.

Childnet polling cited in the Govt response shows that where children use VPNs it is most commonly to protect their privacy rather than to circumvent (CP 1643, p.23), and the response itself records that children typically circumvent age assurance with less sophisticated methods (CP 1643, p.22). Stakeholders warned that restrictions could disproportionately affect vulnerable groups including LGBT+ young people and those experiencing domestic abuse (CP 1643, p.22).

The response recognises that many families run VPNs at home, and its answer is practical guidance telling parents how these tools operate and how to find out whether they are installed (CP 1643, p.24).

- ✦ A 15 year old does homework on the family laptop, which runs a VPN because a parent works from home under company security policy. Circumvention detection flags the household and locks the account, with no remedy clear.
- ✦ A boarding pupil uses the school network, which routes all traffic through a single filtered gateway. Traffic detection cannot distinguish one pupil from another, so the safest compliance decision for the provider is to restrict the whole gateway.

G. Under 18s at Work

The Government response introduces one exemption: chatbots typically used in business or customer service settings will be exempt from the measures, with the definitions of what is in scope to be set out in regulations in due course (CP 1643, p.21). That exemption however does not resolve the position of a 16 or 17 year old employee using a workplace tool. Whether an AI coding assistant, or the assistant embedded in an employer's office software, is typically used in a business setting and therefore exempt, or is a general purpose chatbot and therefore in scope, is exactly the question the drafting leaves open. The employer cannot plan against it, and the compliance-safe answer is to strip the tool or to age-assure the workforce.

The difficulty is that the exemption appears to be based on the setting in which a tool is used rather than on what the tool is, nor is it clear how the exemption would be applied in practice. An apprentice does not stop being under 18 at the office door, and the software does not stop being business software because they are the one using it.

A definition based on the chatbot function would be both simpler to apply and closer to the risk the Government has identified. Tools whose purpose is productivity, coding, drafting, scheduling, translation, customer service, do not simulate a relationship and do not generate the emotional attachment the measures exist to address. Placing them outside the chatbot rules in their own right, rather than by inference from the setting, would give employers something they can actually comply with and would remove use cases the Government does not mean to include in scope.

None of the Government documents addresses employment expressly. 16 and 17 year olds are apprentices, trainees and employees, chatbot use is now ordinary in many workplaces, and the rules follow the person, not the setting. The Government's own case for digital skills makes the problem. The consultation document said children will need digital skills for "a job that hasn't yet been invented" (March 2026), the June statement called learning to use chatbots an increasingly important digital skill for young people (Progress statement, p.30), and the response commits to an enhanced computing curriculum covering AI, data science and technological bias, with first teaching from September 2028 (CP 1643, p.36).

The cost of leaving this unresolved is considerable. Whilst a large employer can seek advice, segment its licences and absorb the compliance work, a firm of fifteen people cannot, and its safest course is to withhold the tool from its youngest employee. That would be at odds with other areas of Government policy, including the purposes of incentivising more apprenticeships across the country.

- ❖ A 17 year old apprentice software developer uses an AI coding assistant, standard practice across the industry. Mandatory breaks interrupt paid work that adult colleagues at the next desk continue without.
- ❖ A 16 year old on a T-level industry placement is asked to draft customer emails with the firm's AI assistant. The employer must first age-assure its own staff member and apply child settings to a workplace tool.
- ❖ A 17 year old marketing apprentice at a small business drafts the week's social posts with a chatbot. The break rule applies at work as at home, and the employer is left planning its working week around a child protection measure..
- ❖ A 17 year old admin apprentice uses the assistant embedded in the firm's office software. Whether the business exemption reaches it is undefined.
- ❖ A school leaver in a first job is the only member of the team timed out of the firm's AI tools. The result is a two-tier workforce in a labour market where the Government is simultaneously funding apprenticeships in exactly these skills.

H. Restrictions on Sexualised and Intimate Features

The June and July positions differ in this area. The July Government response restates the June package as preventing under 18s from accessing chatbot services that primarily offer sexualised content, and barring other chatbot services from offering children features that enable sexually explicit or sexual role-playing content (CP 1643, p.7). The June statement went wider, committing that "similar intimate functionalities will be restricted for under 18s on AI chatbots more widely" (Progress statement). "Intimate" is not defined in either Government document.

Whilst the July wording is directed at sexual content and sexual roleplay, the June wording is not. Providers configure against risk rather than against final text, so the wider reading is likely to shape behaviour before any regulation is made. The response separately records stakeholder concern that restrictions in this area could disproportionately affect vulnerable groups, including LGBT+ young people and those experiencing domestic abuse (CP 1643, p.22).

- ❖ A 17 year old asks a chatbot to explain consent for a PSHE assignment. The topic trips an intimacy filter, even though the same material sits on the statutory RSHE curriculum from September 2026 that the same document relies on (CP 1643, p.36).
- ❖ A sixth-former asks for feedback on the romantic subplot in their creative writing coursework. Refused, despite creativity appearing on the Government's own list of benefits (CP 1643, p.18).
- ❖ A GCSE English student asks about love poetry in the set anthology. An over-cautious filter declines to engage with the exam texts.

Conclusion

Every difficulty raised in this paper returns to one unresolved question: whether a chatbot is defined by its interface or by its purpose. Defined by interface, the measures reach assistive technology, workplace software and agentic tools that have little in common with the use cases that are concerning policy-makers. Defined by purpose, they reach the services the Government has made a case about. That choice has yet fully not been made, and until it is, every other parameter is being set against an unknown scope.

Five of those parameters remain open on the Government's own account:

- ❖ break intervals and duration, referred to the Expert Panel (CP 1643, p.19)
- ❖ the intimate functionalities definitions, stated more widely in June than in July and unresolved between them
- ❖ the harmful, inaccurate or unverified mental health advice standard (CP 1643, p.20)
- ❖ Definitions for the business and customer service exemption (CP 1643, p.21)
- ❖ the age assurance required to enforce any of it, for which the only commissioned work is scoped to the 16 threshold and not the 18 threshold the chatbot measures use (CP 1643, p.23).

The parameters are unsettled by the Government's own admission, and the timetable is short: first regulations before Parliament by end 2026, further regulations within a year of publication (CP 1643, p.4-5), with in-scope definitions to be set out in regulations in due course (CP 1643, p.21). Definitions of this consequence should not be made in guidance once policy decisions are made, as service providers will need to configure against uncertainty long before the guidance arrives.

It is worth noting recent research by the Centre for the Governance of AI who found that 7% of frontier model releases were delayed or withheld from the UK relative to the United States, with regulatory uncertainty rather than any single rule the primary driver, and that delay rates fall as a regulatory environment settles. The Government is rightly pursuing an AI adoption agenda and this set of measures on parallel tracks, and there is a risk that these potential measures inadvertently undermine the wider AI objectives.

A better outcome is available to policy-makers without abandoning the objective. It would distinguish services designed to simulate an ongoing relationship from general purpose assistants, assistive tools and agentic products, and scope the policy interventions accordingly. This would mean reserve mandatory breaks for the narrow category the evidence supports while relying on reminders, prompts, time limits and parental controls elsewhere; confirm a clear definition of chatbots, intimate functionality, unverified mental health advice and the business exemption before regulations are laid; commission age assurance work at the 18 threshold these measures actually use; and test the drafting against circumvention at the design stage rather than in a review three years later. None of that weakens protection for the children the Government is concerned about but rather it would respond to the available evidence.

Appendix: Whether the Measures Would Work

A rule that is straightforward to avoid does not protect the young people it is written for. Each measure as announced has potential circumvention routes around them, and each route carries its own policy consequence that Government decisions on what comes next should seek to anticipate. Indeed, the Government's own evidence base supports several of them.

1. **Service-hopping to beat time limits.** If breaks are enforced per service, a child timed out of one chatbot moves to a second, then a third. General purpose chatbots are free, numerous and interchangeable for homework use. The break rule then binds only children who use a single service, and the compliant provider loses the user to a competitor mid-session. The only fix would be cross-service usage tracking which carries privacy and wider implications.
2. **Age misdeclaration at sign-up.** The Government's own cited evidence puts this first: Internet Matters found that 32% of children admitted bypassing age checks, and the most common method was entering a fake birthdate, at 13% of 9 to 16 year olds (CP 1643, p.23). Without verification at account creation the breaks and feature restrictions never attach. Yet with verification, the June promise that general purpose models would not be age gated would no longer hold.
3. **Borrowed adult sessions.** The second most common method in the same evidence, at 9% of 9 to 16 year olds (CP 1643, p.23). The family laptop, a parent's logged-in account, an older sibling's phone. A 15 year old in an adult session receives no break, no filter and no protection at all.
4. **Local and open source models.** A capable model downloaded onto a laptop runs with no age checks, no breaks, no content filters and no oversight. Regulation that binds only hosted UK-facing services shifts exactly these users beyond its reach.
5. **Chatbots embedded in other products.** The Government concedes the scope problem in principle, acknowledging that some chatbots fall outside the Online Safety Act because of the way they function and committing to close that loophole (CP 1643, p.19). It does not resolve it for these measures. Assistants inside search engines, office software, games and social platforms may or may not be chatbot services within the regulations, but the definitions are deferred (CP 1643, p.21). Children will find whichever surface the rules miss, and providers will face pressure to disable UK features rather than litigate the boundary.
6. **Reframing to defeat filters.** Children rephrase blocked questions as fiction, homework, journalism or questions about a friend. Filters tightened in response then overblock further. The Government also needs to be clear on realistic expectations: chatbots may exclude some usage by policy and make proportionate efforts to prevent use outside those terms and conditions, but that will not mean teenagers never use platforms outside the proscribed limits.

Ofcom's research and report into circumvention is not due until spring 2028 (CP 1643, p.9), by which time the chatbot regulations will have potentially been made. The approach to compliance and circumvention should be a design test for the regulations rather than a review question after them, and on current drafting none of the measures passes it.

Sources

- ❖ [Growing Up in the Online World: Government response](#), DSIT, CP 1643, July 2026. Cited as CP 1643.
- ❖ [Growing up in the online world: progress statement](#), DSIT, June 2026, cited as Progress statement with page references to the print edition
- ❖ [Written Ministerial Statement HCWS254](#), 15 July 2026, cited as WMS
- ❖ [The Secretary of State's oral statement](#), 15 June 2026
- ❖ [Delays to Frontier AI in the EU and UK](#), Centre for the Governance of AI, 2026