fipr foundation for information policy research

To whom it may concern,

This is the response from the Foundation for Information Policy Research (FIPR) to the government's *Growing Up in the Online World* consultation. FIPR is a research foundation which produces expert technical, legal, and policy evidence on issues at the intersection of technology and society, drawing on the expertise of our members and invited experts. We welcome the fact that the government has insisted this consultation should run its course following Royal Assent for the Children's Wellbeing and Schools Act 2026 S.70-72 on 29 April, and value this opportunity to contribute our views.

The 2026 Act amends primary legislation, notably the Online Safety Act S. 214, as well as the Data Protection Act 2018 and UK GDPR. It grants sweeping powers to the Secretary of State over a high-risk technology, without Parliamentary scrutiny and with little provision for guardrails against violations of privacy, freedom of expression, or security breaches. It assumes systems will be accurate and provides no safeguards against error.

It further requires age-based restrictions on social media for under 18s to be in place by July 2027, mandating Internet services to verify that their users meet a minimum age requirement. The government's position as expressed in Parliament by the Minister, Olivia Bailey, was clear that "*any outcome"* of this consultation will impose some form of restriction "*for children under 16"*. We are responding on this basis.

The timeline allows just 15 months for the government to lay regulations. We have a number of concerns in light of the technological and legal complexities, and we urge the government to seek - throughout the process - the views of a broad multi-stakeholder community, including a range of experts from technical and other fields, and including young people themselves.

We are particularly concerned about the proposed age-based restrictions in the Act[1]. We urge the government to consider and obtain evidence as to (1) whether the proposed restrictions or their implementation will meaningfully achieve the intended effect of reducing harm, and (2) whether these restrictions and their implementation may bring additional harms. We would also note that all 'hard' technical assurance systems will apply to adults as well as children.

Many of the questions posed in this consultation make assumptions that viable technologies exist for age verification for deployment by end-users including children and separately by

[1] This includes the potential for widening the scope of providers and the amount of discretion granted to them, potentially to include all those in scope of the Online Safety Act, and further to include network-level providers.

service providers. Any measures intended for deployment by end-users, parents/guardians, or educational establishments will need to be easy to use, prevent evasion, and low-cost. The questions also assume that there are ready means of determining whether breach of any legal requirements of social media companies has occurred and/or that measures were inadequate. The law will need to provide clarity and certainty of expectations. Any Ofcom enforcement mechanism will require reliable evidence. Any measures will need to be able to discriminate at age range borders - for example, to distinguish a 15-year-old from a 16-year-old. We set out details below. Although we understand the concerns that prompted many of the questions in the consultation, to the extent that technology-based solutions cannot meet these aims, policy emphases will need to steer towards alternative, more achievable routes. We believe it would be helpful if the government and Ofcom explained that no sets of laws and regulations can provide complete protection from online harm.

In this response, we assess the effectiveness of age assurance and circumvention technology with regard to the aims of this law, and make a number of recommendations. We discuss a number of issues facing 'effective' age assurance[2] (as sought by the consultation), and additionally provide evidence relating to the technical feasibility of implementing these systems, their negative effects, and the risks and harms posed by circumvention. As we argue and evidence below, many of the proposed implementations of age restriction have limited positive effects in reducing harm to children, while causing significant additional harms, especially to the most vulnerable children and adults. However, we make a number of positive recommendations as to what we consider the best ways forward.

Technical implementations of age verification systems and 'effective' age verification

A range of technical implementation options are proposed for age-verification, each with its own benefits and challenges, and with a greater or lesser reliance on 'strong' technological assurances versus technosocial approaches that go beyond purely technical aspects in their design. The age verification debate rightly foregrounds particular harms to children from exposure to harmful content, harms from 'addictive' app design, risks to children's privacy, the security of their data, and risks posed by those who seek to exploit children directly. As we argue here, many of the more technically-focused approaches to age verification either do not effectively mitigate these risks, or in fact increase children's exposure to them.

Direct assurance systems

Direct assurance systems are used on a range of existing services, involving either themselves collecting data (such as biometric data or credit card information) in order to assess the age of those opening an account, or by relying on a third-party shared service that collects the data and uses it to provide authentication across multiple sites. This poses a number of

2 As defined by the OSA S.12[6] and S.161 [5] and underscored by Ofcom in its Protection of Children Code of Practice, namely, as meaning age verification and age estimation that correctly determines whether or not a particular user is a child.

serious *security* and *privacy* issues to both children and adults (as both are forced to use these systems). This approach requires trust in the site and assurer not to misuse, insecurely store, or resell the sensitive biometric and financial data provided (the risk of a ‘data grab’ by websites). In 2018, Facebook came under fire for using phone numbers provided for account verification for targeted advertising and for finding user accounts.

This approach also links the use of a site to a verified identity, which, if compromised, enables the collection, exploitation, and even monetisation of sensitive data, including through blackmail, private surveillance, and serious crime. Further, approaches that ‘detect’ age face well-understood challenges in dealing with people close to an age boundary (such as 17-year-olds) and tend to perform poorly with users from minority ethnic, disabled, LGBTQ, and other structurally disadvantaged communities. This bias is built into systems largely trained on the data from people who do not present to digital systems the way that these communities do. A system that in practice sets a higher bar for access to social media or other sites for minority ethnic or female users because it under-estimates their age causes exclusion from democratic participation for these already-marginalised groups.

Further, these systems normalise repeated age checking across the Internet, helping hostile sites to use it to steal biometric data or credit card information. In practice, they are also generally easy to circumvent; media report rising use of age manipulation filters and other widely available technologies.

Anonymous credentials systems

Anonymous credentials systems use cryptographic methods to provide proof of eligibility. A trusted issuer checks the conditions of eligibility (for example, age) and issues a credential. This age verification (or other) credential can be used on any site or service that trusts the issuer to prove eligibility. However, it communicates only that the person is eligible and nothing else. This 'data minimisation' is important in limiting the risks of trusting the issuer. These credentials also have the property of non-linkability. Because they change each time they are used, the service cannot link multiple visits to the same site or visits to many sites to into a pattern of activity. Even the assurer who issued the credential cannot tell the identity of the user of the credential when they present it. Anonymous credentials therefore provide very strong technical assurances for security and privacy and are often suggested in the context of digital credentials ‘wallet’ schemes.

However, anonymous credentials systems create a number of other serious issues with respect to implementation. First, they heavily concentrate trust in the issuer; all users and services need to trust the institution which issues the credential, who is in a position to provide credentials fraudulently or deny them to eligible users. Widespread adoption of these systems as a condition for accessing services and platforms causes serious issues of digital exclusion for those without formal identification, for those with forms of identification the issuer does not accept, or accepts only in a different format, and for those who do not trust the public or private institution that issues the credential.

Further issues come from the users' ability to transfer their credentials, and hence access, to others, either by transferring the credential directly, or by giving another user access to a verified account. Research on cybercrime forums consistently shows a lively trade in fake and stolen ID documents and credit cards, credentials, and verified accounts. The solutions that do exist, such as tying the credentials to specific hardware like a mobile phone or smartcard are flawed. They very quickly extend the range of trusted actors and still do not prevent hardware transfer. In addition, usability remains a serious challenge for these systems; it is very hard to explain to many people how they work and why they are trustworthy. The cryptographic properties that make them secure such as non-linkability and data minimisation appear to be impossible to most people when explained, and they struggle to find adoption beyond small enthusiast user bases.

Finally, these systems rely on the site to comply with the requirement to check credentials, therefore causing *displacement*. Sites that want to comply with the law are likely to implement strong, even over-strong age verification. The result will lead from a scenario in which:

- Most curious or accidental underage users attempt to access a site and are prevented.
- Persistent, motivated underage users get through the restrictions to access adult content on sites that generally comply with the law, practice mandated content moderation and remove illegal content

To a scenario in which:

- Most curious/accidental under-age users attempt to access a site and are prevented.
- Persistent, motivated underage users will move to openly-accessible sites that care about complying with UK law, and which feature far more harmful or illegal content than the ‘compliant’ adult sites
- Especially persistent users will use riskier and more harmful circumvention strategies (such as borrowing a device) to access age-gated sites which then gives the device's owner access to sensitive data about them and creates risks of blackmail and coercion.

Because of these issues, *anonymous credentials* systems are not currently widely used.

Tagging and blocking

A more socially-embedded approach, which relies less on complex technical architectures, is **tagging and blocking.** In this system, the providers of sites and services declare, through a system of content tags in the site's HTML source code, the kind of content they host, and the audience they consider appropriate. Parents or trusted adults input the child’s age and other content settings on their device, which compares these details to the tags, and block or allow access on this basis. This has the benefit of being technologically simple and easy for users to

understand. It also has the benefit of giving parents and other trusted figures in the child's life agency. It roots trust in those who know the child, are responsible for them, and have in-person contact with them. Under tagging and blocking systems:

- Most curious or accidental under-age users attempt to access a site and are prevented.
- Persistent, motivated under-age users bypass the restrictions to access adult content on a site that generally complies with the law, practices mandated content moderation and removes illegal content.

Tagging and blocking has the further advantage that it is already well understood by the UK public via a clear, transferable mental model with wide public acceptance - the British Board of Film Classification age ratings and content classifications. Additionally, by allowing safer circumvention by extremely-motivated young people, it provides a safety valve for situations in which children need to seek help and advice *against* adults with authority (such as parents and relatives) who are harming them, or in scenarios such as those faced by LGBTQ children whose parents unduly restrict their ability to access information and community. It also reflects the nature of online content, which in the UK is regulated as speech, media, and social interaction rather than as a 'harmful substance' like tobacco.

Child-safe phones

Child-safe phones can combine many of these methods and may be an element of a layered approach to age verification. They can be designed in a variety of ways, and under different legal and policy regimes. For example, they can take the form of specialist product lines made for children, or can be a 'child mode' that is mandated on every phone and turned on by default unless the user verifies their age. These two examples have radically different consequences for security, privacy, fundamental rights, and unintended harms. The method of age verification used on the phone matters as well. For example, age assurance on child-safe phones can take the form of an anonymous credential issued by a trusted issuer as described above. Alternatively, user age (or under-18 status) can be input by the user's parent. Work by Rebel Tech Alliance in its guidance on online safety for parents recommends privacy- and security-protective phones and apps with effective use of content controls, especially around minimising contact with 'addictive' design features.

Dynamic assurance technologies

Finally, a range of **dynamic assurance technologies** are in the early ages of development. Developers of these systems argue that age assurance should extend beyond a 'one-time' check reliant on a single source of data into using multiple methods. These include stylometric analysis of typed speech, behavioural data, location data, and administrative data to assess the age of a user, or the use of a variety of proxies for vulnerability to dynamically identify different kinds of digital risks. Privacy preservation measures for dynamic systems, where the burden is shared between systems operating at the service, device, and network level are a next stage of innovation in this field, as suggested by Dr Chelsea Jarvie in the 'layers of care' approach.

These systems have the advantage of not conflating ‘age’ and ‘vulnerability’ – they are ‘inference’ rather than verification technologies. That also means they would not be considered sufficient under the current policy regime. While there are potentially secure and privacy-preserving ways to design such systems, they pose a range of serious practical issues (for example, explaining how they work to users). As with all algorithmic systems, there are potentially serious privacy and equity issues. For example, the system may consistently assess particular social groups to be more vulnerable and thereby disproportionately exclude them.

Implementation issues

To summarise, each approach to designing age verification systems has a distinctive balance of social and technical elements. It is undoubtedly tempting to prioritise the strongest possible technical assurances in age verification systems - and we agree that strong assurances on privacy and security for children and adult users should be at the heart of any adopted system. However, we argue that over-reliance on technical design elements in order to impede circumventing these systems as much as possible has the effect of intensifying the very harms these systems are meant to prevent. We now consider a range of circumvention strategies and associated harms.

Circumventing age verification and restriction systems

It is important to make a distinction between systems designed to prevent ‘accidental’ or ‘curious’ access to restricted content and those intended to frustrate a highly-motivated user bent on circumventing in-built restrictions. We argue that almost no age verification system, however technically secure, can frustrate the most motivated attacker (even a child). This is partly due to technical limitations, but is also a consequence of embedding these systems in already-existing social contexts, which produce a range of circumvention methods. As we discuss below, while it's essential to attempt to improve systems' privacy and security properties, attempting to secure a system against all possible circumventions or risks itself creates significant usability issues, risks, and threats. This issue has been well-documented over decades of information security scholarship.

Virtual private networks (VPNs) are one way of circumventing age-based restrictions as outlined in a Joint Statement of security and privacy academic researchers. Many approaches besides VPNs offer children the ability to circumvent age verification and other country-specific rules. The potential for unintended harm inherent in these other circumvention methods as a consequence of ‘strong’ forms of age verification, accruing from the fact that many of them centralise extremely sensitive information in the hands of actors who can exploit it to harmful ends.

Privacy Enhancing Technologies

There are other technologies which, like VPNs, allow falsifying the user's geolocation. These allow non-gated access to sites which only implement age verification for UK users. One such service is Tor, a global network of servers run by volunteers which provides its users with strong anonymity and privacy protections as well as extremely strong security and anti-censorship properties. By design, neither Tor network nodes, nor the destination website, nor the user's Internet Service Provider know both the source and destination of the signals, nor can they block these requests. Unlike the methods we discuss below, Tor does not centralise trust in a single provider such as a VPN company or the owner of a shared phone. The Tor Browser can be downloaded for free from the Tor Project website.

Blocking or banning the Tor network (as proposed by many previous UK governments) is theoretically possible, as the network publishes a list of these relays, which Internet Service Providers could feasibly block. However, this would have broader negative effects. Tor is widely used by law enforcement and security professionals for its security and anonymity properties, and it underpins journalistic provisions by organisations such as the BBC and the Guardian to protect whistleblowers. Many civil society and third sector groups also use it for self-protection when working in dangerous environments. Other circumvention services such as Snowflake, a collaborative effort where Internet users provide access points to the network for others provide access in jurisdictions where Tor nodes are blocked. Blocking the Tor website, as many 'school-friendly' ISPs' do, would prevent downloading the Tor browser for at least some children. However the Tor browser is also available as an app on app stores for both Apple and Android phones. Over many years, authoritarian regimes' attempts to block access to Tor have been largely unsuccessful. Tor is a very strong circumvention tool which is free, fairly easy to use, and effective against all proposed forms of country-specific age verification. Its existence makes blocking or age-gating VPNs an essentially pointless and harmful exercise.

Direct circumvention

A different range of widely-reported approaches allow direct circumvention of age checks by websites, apps, and services. Age verification also creates demand for **stolen data markets**, which provide the means for circumvention. These underground markets are often centred on forums or Telegram communities. Personal documents (photographs and scans of official documentation used for identification) are readily available for sale on these markets (see for example Marjanov et al., 2026). Furthermore, when age verification datasets are inevitably breached (for example, the 2025 Discord data breach), there is a ready market for distribution. Since some systems currently use credit cards as a proxy for age, we also note the existence of a mature, extensive, and well-developed online market for stolen credit cards, which can be purchased extremely cheaply on many underground sites, including those on the clear web.

Children have already proven adept at finding methods for 'tricking' age detection software in the months since the Online Safety Act restrictions took effect. One much-reported workaround involves using tools from the popular videogame *Death Stranding* to generate realistic facial images with dynamic expressions which fooled Discord's age verification systems. Age-changing filters and other video editing tools are widely available in mainstream camera and

social media apps. While newer, more robust versions of age check software will add features like filter detection and therefore be harder to circumvent, adopting this approach sets up an arms race between services and circumvention software developers. This again risks driving young people to illicit markets where these technologies are sold rather than sticking with the filters available on mainstream apps.

Device sharing

A further concern involves **device sharing**. This takes many forms: children may purchasing an age-verified device in the clandestine market, use an older sibling or parent's device, share a device with a community of friends at school, or rent a device from an older fellow student or neighbour. Each approach has significant associated harms, especially as they provide the device's original owner with access to the child user's sensitive data and details of what they have done with the device, which likely involves accessing adult content. This creates a risk of coercive control, bullying, blackmail, and further exploitation. Given the dynamics of school bullying and gender-based violence, and increasing concerns about the rising influence of the manosphere in UK teen culture, these risks and harms are likely to fall particularly on young women. Marginalised or otherwise vulnerable groups - such as LGBT children - will be particularly at risk.

Purchasing credentials and account trading

The online ecosystem of account reselling and trading services already exists and is readily accessible by children. In this model, accounts for over-18 sites and services are created with valid credentials and then stolen, resold, or simply given to under-18s. Existing markets for reselling accounts on Steam, video games, Amazon, social media, pornography sites, and a wide range of others are lucrative and easily accessible. To give only three of many examples, a wide range of accounts for 16+ and 18+ services are available for purchase on Hack Forums, which is listed by Google search. The Accszone website (also listed on Google Search) sells verified social media accounts, with current prices for a single Facebook or Instagram account starting at around US $0.80. There is also a large market for verified accounts for adult services on Telegram channels. The main anti-circumvention method for account sharing or trading is device-locking, - a set of measures which allow access only from the particular device detected by the service. However, locking would place leave adult Internet users unable to access most Internet services except on specific devices, and burden websites and service providers with the need to implement hardware device-locking for their accounts. The need for multiple-device access for almost all sites (both for convenience and integration of services) would provide easy means of circumvention.

Age-based restrictions and the Children's Wellbeing and Schools Act, 2026 S.70-72

We close by setting out a number of concerns with respect to the Children's Wellbeing and Schools Act, 2026 S.70-72. First and foremost is the impact on both children and adult users of social media, and the methods of data collection by age verification systems. This law is

claimed to regulate large companies, but in fact regulates individuals. Individuals must verify their age in exchange for access and account holding rights, and Internet services must enforce it. By default, age verification checks all adult users as well as child users, and thus all adult social media users will have to show ID to make accounts and post or view content. These systems are incapable of singling out children.

The implications have been overlooked in the legislative process, notably in the Parliamentary debates on the Children's Wellbeing and Schools Act 2026 s.70-72. We are concerned that this policy is being put in place based on the assumption that it will operate just like offline age-based restrictions such as alcohol purchasing. Instead, replicating off-line age restrictions in online spaces presents multiple challenges, according to research from the Turing Institute. Ultimately, the comparison is flawed because showing ID in a shop to prove age does not create a data trail of every communication and location change.

Age assurance raises fundamental privacy challenges, as outlined in this IETF Draft on Age Verification Architecture. Age assurance providers will collect data, not only at sign up but in every interaction, including contextual and behavioural data as described in a DW report about Meta's plan using AI to analyse social media profiles. Academic research suggests that behavioural profiling raises privacy and security concerns and is legally problematic. A particularly egregious possibility is that the data will be used for AI training, something to which no child or adult can meaningfully consent and which conflicts with data minimization requirements. Further warnings about the negative impact of age assurance systems that rely on facial recognition are outlined in a joint UK-EU paper. We are concerned that age-restrictions create a barrier with which every user will have to comply before they can participate in online life or communicate with others. In other words, individuals are being asked to give up their data merely in order to speak. This raises serious questions around compliance with Online Safety Act, Article 22 and with the right to freedom of expression as defined in Article 236, and we urge the government to ensure this precious right is protected.

Critically, age-checking technologies do not tackle parents' central concerns about the viral dissemination of toxic and harmful content. That problem will remain unless the underlying harms of social media platforms are tackled by both government and Ofcom. Our concern is that these measures will entail AI-driven content moderation, which poses additional risks to freedom of expression and privacy rights.

Recommendations

R1 - We urge the government to seriously consider the potential harms which specific implementations of age verification regimes may cause. We argue that requiring ‘strong’ technical implementations of age verification systems for access to Internet services will not significantly prevent harms to children, and will expose both children and adults to the very risks that this policy aims to mitigate.

R2 - Approaches requiring 'strong' technical restrictions on access to Internet services for all UK users are not a matter of mandating a single technology, but in practice will entail developing a new infrastructure of age verification that sits atop the Internet. If left to private providers, this infrastructure will risk serious issues with security, and has the potential to aggregate sensitive user data in the hands of profit-making companies. Recent serious scandals around the management of UK infrastructure - from Thames Water to Grenfell Tower - illustrate the harms and perverse incentives which emerge when industry is asked to ‘mark its own homework’. Social media providers' procedures to protect and secure user data for age verification purposes should be subject to strong regulatory oversight. Public provision of age verification infrastructure has its own issues, from excluding particular groups to problems of explainability, trust, and legitimacy.

R3 - It is clear that age restrictions and technical solutions alone will not resolve – and may worsen - the central problems they are meant to tackle. Placing the regulatory burden on users rather than on the major platforms, their design, and business models will not address the sources of these harms. Any age verification system should integrate the involvement of individuals and communities with responsibility for children, and should facilitate, rather than shut down, conversations with children about their Internet use. Much of the policy debate has been conducted ‘over the heads’ of children, with little consideration of their views on safety and the Internet. Implementations should minimise security and privacy risks to all users. Regulation should aim to protect the right to freedom of expression.

R4 - It is our view that of the measures we describe *tagging and blocking systems*, in which sites flag their content and recommended age limit, and the child’s age and content restrictions are set by parents on the device itself, are potentially a positive middle ground. These approaches will prevent curious or accidental access to risky spaces and will minimize the risk of pushing children to the most harmful methods of circumvention. They will also minimise the privacy and security harms to adult and child users. ‘Child-safe’ secure phones for younger users may have a part to play in this, depending on how protections are implemented.

Yours,

Dr Ben Collier
Chair, Foundation for Information Policy Research

Dr Monica Horten,
Vice-Chair, Foundation for Information Policy Research

We would like to thank the FIPR members who contributed to the writing of this document, and particularly expert members consulted for technical advice on specific points, including Professor Alice Hutchings on issues with circumvention and displacement, Prof Steven Murdoch regarding anonymous credentials systems, and Dr Daniel Thomas on tagging and blocking designs, among others. We also thank the experts who attended the FIPR Round Table on Age Verification for their participation, and for the discussions which informed this response.