ORG OPEN RIGHTS GROUP EFF

UK ONLINE SAFETY BILL WILL MANDATE DANGEROUS AGE VERIFICATION FOR MUCH OF THE WEB

September 2023

Under new age verification rules in the UK's massive Online Safety Bill, all internet platforms with UK users will have to stop minors from accessing 'harmful' content, as defined by the UK Parliament. This will affect adult websites, but also user-to-user services – basically any site, platform, or app that allows user-generated content that could be accessed by young people. To prevent minors from accessing 'harmful' content, sites will have to verify the age of visitors, either by asking for government-issued documents or using biometric data, such as face scans, to estimate their age.

This will result in an enormous shift in the availability of information online, and pose a serious threat to the privacy of UK internet users. It will make it much more difficult for all users to access content privately and anonymously, and it will make many of the most popular websites and platforms liable if they do not block, or heavily filter, content for anyone who does not verify their age. This is in addition to the dangers the Bill poses to encryption.

The details of the law's implementation have been left to the UK's regulation agency, the Office of Communications (Ofcom), but the Bill is vague on the details of this. Social media and other sites, where users regularly engage with each other's content, will have to determine the risk of minors using their site, and block their access to any content that the government has described as 'harmful'. Platforms like Facebook and TikTok, and even community-based sites like Wikipedia, will have to choose between conducting age checks on all users – a potentially expensive, and privacy-invasive process – or sanitising their entire sites. That's why Wikimedia has come out strongly against the Bill, writing that in its "attempt to weed out the *worst* parts of the internet, the Online Safety Bill actually jeopardises the *best* parts of the internet".

Providers of pornographic or 'adult only' services will, of course, have no choice except to impose age verification to exactly identify the age of the user and not allow under-age users onto their site at all.

The government's list of content that is harmful for children includes violent content and content relating to eating disorders, suicide and even animals fighting. This list will be enshrined in law, but contains no further definition, leaving it open to misinterpretation. It is impossible for a large platform to make case-by-case decisions about which content is harmful. For example, a post which describes a person overcoming such a disorder, a post describing necessary health information and advice about the topic, and a post explaining how much weight a person lost as a result of an eating disorder could all be described as eating disorder-related content that is 'harmful'. As a result, services will be forced to over censor to ensure young people – and possibly, all users, if they aren't sure which users are minors – don't encounter any content on these topics at all. Site operators will undoubtedly be liable for errors, and many sites will require over-zealous moderation to ensure they are complying, resulting in lawful and harmless content being censored.

This leaves only a few options for platforms, services, and apps with UK users, and all of them lead to a less open, less functional, and less free Internet. Platforms will face criminal penalties for failing to comply and may choose to block young people – including those as old as seventeen – entirely. They may filter and moderate enormous amounts of content to allow young people on the site without age verification. They may filter and moderate enormous amounts of content for young people only, while allowing age-verified users access to all content. Or, they could exclude UK users entirely, rather than risk liability and the cost of expensive and untried age estimation systems and content moderation.

Whilst the policy aim is well-intentioned, the result will be dangerous. The requirement to age-gate will trump the balancing of rights. It risks a disproportionate interference with children's and adult's right to access information, and their freedom of expression rights.

Which sites will be affected?

The Bill primarily covers two types of sites: web services that solely exist to publish and sell access to pornographic content, and user-to-user services which allow users to post their own content. These platforms may carry limited amounts of pornographic or 'harmful' content – because user-generated content is impossible to moderate at scale – but clearly that is not their primary purpose.

Pornographic websites will have to prevent under 18s obtaining any site access at all. Social media platforms and other sites that contain user-generated content, on the other hand, will have to assess the risk of children using their service, and the risk of content defined as harmful to children being on their site. They will have to block children from being able to access content defined as harmful. This includes pornography, but the full list encompasses a much wider range of content (see below).

Adult and pornographic websites

Pornography websites that have UK users, or target UK users, will be required to use age verification to ensure that children are not able to encounter their content. Age verification is, essentially, identity verification, which makes it effectively impossible to browse pornographic sites anonymously, and creates the risk of data breaches and the potential for data to be collected and potentially shared or sold. Data protection laws apply, although little guidance exists in the Bill about compliance. Ofcom is responsible for determining the measures and policies sites should implement, and the principles that will be applied to determine compliance [S.83]. The Bill does explain that sites should "have regard for the importance of protecting UK users from a breach of any statutory provision or rule of law concerning privacy that is relevant to the use of operation" of the service. Privacy should be paramount in a bill like this, but for now, how exactly that will happen has been left to Ofcom.

Social media platforms

Social media platforms which allow minor users will be mandated to deploy technical solutions to check the age of users before serving content. This is clear from S.12, 'Safety Duties protecting Children'.

Online platforms must prevent children of any age encountering "primary priority content harmful to children," [S.12 (3a)] and to "protect children in age groups judged to be at risk of harm from other content that is harmful to children (or from a particular kind of such content) from encountering it by means of the service [S.12 (3,b)]. Platforms also now have to consider how to protect children from “features, functionalities or behaviours enabled or created by the design or operation of the service” [S12 (3,C)].

Platforms will also have to conduct a risk assessment to explain how they will address children of any age and those in age groups judged to be at risk of harm [S.11 (6)]. They are expected to comply using age assurance, age verification or age estimation [S.12(4), 12(6) and S.12(7)]. Age estimation likely involves estimating age based on biometric data – essentially, using an algorithm to scan a photo or video of the user.

What content is covered?

The Bill describes two types of content: primary priority content and priority content. But there’s little relevant distinction in practice. Children must be “prevented” from access to primary priority content, which suggests they must be blocked from accessing it at all times, whereas children should be “protected” from coming across priority content, but the measures required are the same. The Bill does not explain the distinction between “prevent” and “protect” in this context.

“Primary priority content” has been confirmed in the law. The list specifies pornographic content, but also includes content encouraging, promoting or providing instructions for suicide, self-harm (including poisoning) and eating disorders. [S.61] Priority content is anything depicting violence against people or animals (including fictional animals) [S62 (14)], bullying content, abusive content related to a number of protected characteristics, content that promotes dangerous stunts (such as the cinnamon challenge), and content which encourages people to “ingest, inject, inhale or in any other way self-administer” a physically harmful substance, or any substance in quantities which would be harmful [S62.9].

How will age verification work?

Age verification is defined as any measure to verify the exact age of a user. In practice, there are two types of verification. The first, commonly called age verification, usually involves confirming a user matches with government issued identification. The second is age estimation, a measure intended to estimate the age or age range of a user based on their appearance. Self-declaration will not be accepted for compliance purposes. Providers will have to design their services to take account of the needs of children of different ages, and ensure that there are adequate controls over the use of their service by children [S. 7(4)]. They can only conclude that children cannot access their services by implementing age verification in such a way that children cannot normally access the service [S.12].

Compliance will be compulsory unless the terms of service of the platform explicitly prohibit the content that is being addressed.

Providers will have to choose systems that are “highly effective at correctly determining whether or not a particular user is a child” [S12 (6)]. Providers can even be required to distinguish between children of different ages, for the purpose of determining whether they can be permitted to access certain content.

There is no privacy-protective age estimation or verification process currently in existence that functions accurately for all users. France’s National Commission on Informatics and Liberty (CNIL) published a detailed analysis of current age verification and assurance methods. It found that no method has the following three important elements: “sufficiently reliable verification, complete coverage of the population, and respect for the protection of individuals' data and privacy and their security.” In short, every age verification method has significant flaws.

These systems will collect data, particularly biometric data. This carries significant privacy risks, and there is little clarity in the Bill about how websites will be expected to mitigate these risks. It also carries risks of incorrect blocking where children or adults would be locked out of content by an erroneous estimate of their age. This risk is recognised by the inclusion of a requirement for providers to consider complaints by users whose age has been incorrectly estimated [S 32 (5)(D)].

Ofcom could minimise the damage of this Bill, as they are required to produce a code of practice on age assurance. The first principle that Ofcom should adopt is that the age assurance or age verification systems should be effective at correctly identifying the age or age-range of users, and that competition of provider should exist so users with a concern for privacy and security can opt for their chosen provider. The pressure will be on Ofcom to ensure that platforms implement age verification or age assurance, and this will have priority over any balancing of free expression rights. This poses a risk to the fundamental rights of huge numbers of users.

Choices for providers

Overall, there are some foreseeable problems with this entire approach. There is significant risk that young people – who could be seventeen – are banned from large swathes of the web. They may well be banned entirely from some platforms and services.

Alternatively, large swathes of content will be removed for all users, including adults, due to over-moderation by providers operating under a strict liability regime. Those users, whilst they are given an option to complain, may find it difficult to do so.

Providers will have a Hobson's Choice between age-gating at the site level and blocking children, ensuring they stay on the outside, or sanitise their entire site to child level. If they don't want to do either of those, they will be required to do age-gating at content level.

The other option is that providers choose not to serve the UK at all.

Risk assessments

Online platforms must also complete risk assessments – a task that may be difficult, if not impossible, for many services. In addition, they must report how they will address children of any age and those in age groups judged to be at risk of harm [S.11 (6)].

A risk assessment also must determine the number of children who could encounter primary priority content on the service, and there must be a separate assessment for each type of content. The platform must re-work the risk assessment every time they have a system re-design. The first risk assessment must be carried out within three months of the Bill coming into law, and records must be kept of each one.

All of this must be done within the first six months after the Bill gets Royal Assent.

For more information, contact: James Baker, Open Rights Group:
james.baker@openrightsgroup.org