TechScape: How the UK's online safety bill aims to clean up the internet

Publication Date: 2023-10-24

Author: Dan Milmo

Section: Technology

Tags: Internet safety, TechScape newsletter, Social media, Digital media, Internet, Children, Ofcom, Pornography, newsletters

Article URL: https://www.theguardian.com/technology/2023/oct/24/techscape-uk-online-safety-bill-clean-up-internet

Deepfakes, viral online challenges and protecting freedom of expression: the online safety bill sprawls across many corners of the internet and it's about to become official. The much-debated legislation is due to receive royal assent, and therefore become law, imminently. The purpose of the act is to make sure tech firms have the right moderating systems and processes in place to deal with harmful material. "This means a company cannot comply by chance," says Ben Packer, a partner at the law firm Linklaters. "It must have systems and processes in place to, for instance, minimise the length of time for which illegal content is present." Packer adds that if those systems and processes are deemed to be not up to scratch, platforms could be in breach of the act. "Even if a platform has no illegal content present on its service, it could still be in breach of the legislation. Likewise, the mere fact that illegal content does appear on a service does not mean it is necessarily in breach." The act is complicated and, as you will see below, contentious. The duties under the bill The bill is based on three fundamental duties: protecting children; shielding the public from illegal content; and helping adult users avoid harmful – but not illegal – content on the biggest platforms. A company is within the scope of the bill if it distributes user-generated material to other users or is a search engine. So it will cover an array of platforms, from X to Facebook, Instagram, YouTube and TikTok as well as Google. To comply with those child and illegal content duties, companies will have to carry out risk assessments on whether that content appears on their service. All of this will take place under the oversight of communications regulator Ofcom, which will issue guidance on what content to focus on preventing and the systems needed to deal with it. Companies can adhere to Ofcom's codes of practice for dealing with this content or come up with their own methods of dealing with such material. But the bill makes clear that if they follow the code of practice, they won't be in danger of breaching relevant sections of the act. According to the act, they'll be "treated as complying with a relevant duty if the provider takes or uses the measures described in a code of practice". In a nod to concerns about the bill being a "censor's charter", all companies within scope have a duty to protect "users' right to freedom of expression" and their privacy. Protecting children The government messaging about the legislation has focused on protecting children. There are two categories of content harmful to children that tech firms must deal with. The first one is "primary priority content", such as pornography and the promotion of suicide and eating disorders (albeit below the threshold of criminality). If sites allow such content, children must be prevented from encountering it and the act expects age-checking measures to be used for this – measures that either target specific pieces of content or cover a specific section of a platform. The second type is "priority content" such as bullying and posts that encourage children to take part in dangerous stunts or challenges. Children in age groups judged to be at

harm from such content – an area where risk assessments will be key – must be protected from encountering this kind of material. Ofcom will set out the steps for doing this in a code of practice that it will draw up. Illegal content The act lists a number of criminal offences that constitute “priority illegal content”, which means companies within the scope of the bill need to have systems and processes in place to prevent users from encountering such material. Those priority offences include: child sexual abuse material; terrorist content; revenge or extreme pornography; and threats to kill. The act wants this sort of content to be proactively targeted by platforms’ moderation systems and processes. There are new criminal offences in the act. People who use social media posts to encourage self-harm face criminal prosecution under a new offence introduced by the bill that covers England and Wales. The act will also criminalise the sharing of pornographic “deepfakes” – images or videos manipulated to resemble a person. Protecting adults from harmful content This applies to “category one” firms, or platforms with the largest reach and greatest influence over public discourse. So you would expect it to apply to the likes of Facebook, Instagram, Google and TikTok. But after a late change to the bill, it could also apply to niche but risky sites like the rightwing-favoured platform 4chan. Once the list of category one sites platforms is drawn up by the secretary of state, those companies will have to give adult users the ability to avoid certain kinds of content if they wish to do so. The sort of content platforms need to protect users from – if the user wishes – is listed on the act: material related to suicide, self-harm and abuse targeted at protected characteristics under the Equality Act (such as age, race and sex). Platforms need to assess how much of that content is on their platforms and put in place a way of shielding users from that content. So this is not a case of clicking a button on these platforms and all the bad stuff goes away. Platforms will need to offer features that, for example, allow users to avoid abuse. This could be an option to consent to someone following you, or a warning screen appearing for certain content or muting certain words. Many platforms offer these sorts of measures already – but now Ofcom is watching. Further duties for category one firms include protecting journalistic content and “content of democratic importance”. Pornography Sites that offer pornographic content must use age-verification measures (jargon for age checks) or age-estimation technology (where you send a selfie to the platform and it gauges whether you are the age you say you are) to ensure users are over 18. This obviously applies to sites like Pornhub and to user-to-user platforms that allow sexual content, like OnlyFans (which already has an over-18 age policy). Sites that are not dedicated pornography services but still allow sexual content, like X, will be expected to apply age-verification or estimation measures either to specific content or a specific section of their service. Ofcom will provide guidance on what are the most effective forms of age verification or age estimation. End-to-end encryption One of the most controversial elements of the bill is a provision on combating child sexual abuse material (CSAM), which empowers Ofcom to order a messaging service to use “accredited technology” to look for and take down such content. Privacy campaigners and tech firms have warned that the clause poses a fundamental threat to end-to-end encrypted messaging – where only the sender and recipient can read the message – because it could require the scanning of private messages. In a bid to head off a threatened exodus, the government said in September that Ofcom would only be able to intervene if scanning content was “technically feasible” and if the process met minimum standards of privacy and accuracy. Some observers took this as a climbdown but the government has not changed the wording of the bill, so this still feels like a very live issue with the legislation. Criminal liability for tech executives Tech executives face the threat of a two-year jail sentence if they persistently ignore Ofcom enforcement notices telling them they have breached their duty of care to children. The government has stressed to jumpy tech execs that the new offence will not criminalise executives who have “acted in good faith to comply in a proportionate way” with their duties. Senior employees could also be jailed if they hinder an Ofcom investigation or a request for information. In terms of punishments for companies, Ofcom can fine impose fines of £18m or 10% of a company’s global turnover for breaches of the act. In extreme cases, it can also block websites or apps. So what next? This is a complicated and wide-ranging piece of legislation, partly a reflection of the sprawling nature of the field it is trying to regulate. Ofcom – and tech platforms – are going to be busy. If you want to read the complete version of the newsletter please subscribe to receive TechScape in your inbox every Tuesday