

How UK Online Safety Bill May Affect Cos. Around The World

By **Lucy Blake and Karam Jardaneh** (November 18, 2021, 11:01 AM EST)

As people around the world are living increasing portions of their lives online, especially during the COVID-19 pandemic, the triumph of the internet in bringing together people, ideas and opportunities has been brought into ever sharper focus.

However, there is also a growing realization that in addition to being a force for good, content on the internet can cause harm. Online platforms have been used to incite extremism and terrorism; spread fake news, disinformation and misinformation; undermine democratic election processes; encourage self-harm; and groom, abuse and bully people, including children.

Moderation of internet content is an inherently subjective process and a potential minefield. It requires a delicate balance between protecting people from harm without attenuating the right to freedom of expression. With calls for laws to protect the public from online harms, the challenge now falls to legislators and regulators to strike this balance.

The new draft U.K. Online Safety Bill[1] is a ground-breaking piece of legislation which, as currently drafted, will police illegal and harmful content online with the stated aim[2] of making the U.K. "the safest place in the world to go online, and the best place to start and grow a digital business."

Extraterritorial in scope, the bill will apply not only to U.K. registered platforms, but also to any platform around the world providing services to U.K. users. If successful, it is also expected to pave the way for further regulation of online harms around the world. In particular, as the U.S. Congress has over recent years introduced multiple bills aimed at content moderation and the EU has proposed a Digital Services Act targeting illegal content, fraudulent advertising and disinformation, eyes are on the U.K. to see how the bill works in action.

This article provides a high-level practical overview of the bill, providing answers to 10 questions and suggesting prudent steps that impacted platforms should consider taking in anticipation of it coming into force. It also provides a brief insight into the status of proposed legislation in the United States.

1. What is the status of the bill — and when will it come into force?



Lucy Blake



Karam Jardaneh

The bill is currently subject to scrutiny by a joint committee of members of the House of Commons and peers from the House of Lords, who are due to report back on their findings by Dec. 10. U.K. Prime Minister Boris Johnson indicated the bill will be debated in the Commons before Christmas, but other ministers have rowed back on this timetable.[3]

The bill has had a mixed reception in the U.K., with some arguing that it has not gone far enough — for example, in protecting people from racially motivated hate speech online and others arguing that the proposals will result in too great an interference to the right to freedom of expression.

The U.K. Financial Conduct Authority has recently weighed into the debate, echoing calls from parliamentary committees that the bill should include safeguards against more forms of online financial scams, in particular fraudulent advertising. The U.K. Law Commission[4] has also proposed that various additional criminal offenses should be enacted to make companies criminally liable for serious harms resulting from online abuse.

Some commentators have gone even further, arguing that provisions should be added to make individual senior managers criminally liable for failing to prevent wrongdoing.

The bill has however been criticized for a lack of precision, particularly in defining what constitutes harmful content. As currently drafted, much is left to the discretion of the secretary of state and clear guidance will be needed to help service providers navigate the new legislation. The current bill also prioritizes the obligation to remove harmful content over the obligation to regard users' freedom of expression, leading to concerns regarding the impact on free speech.

Assuming the bill progresses through Parliament either before Christmas or early in the New Year, it is anticipated that it will come into force at some point in 2022.

2. Who will be subject to the bill?

As currently drafted, the bill will apply to:

- User-to-user service providers that allow users to generate, upload or share content — e.g. social media companies; and,
- Search service providers — e.g. search engines.

Various service providers are expressly exempted from the bill, including, for example, email service providers, short message service and multimedia messaging service providers, internal business service providers, live aural services providers, public bodies and various services with more limited functionality.

There have also been calls to add provisions making senior managers criminally liable for failure to prevent online harms. Currently, the bill contains a deferred provision to make senior managers criminally liable for failure to comply with requests for information from Ofcom, but not for failure to prevent wrongdoing.

The provision on the failure to comply with information requests has been introduced as a deferred power. When the bill was published, the U.K. Department of Culture, Media and Sport said in a press release[5] that the deferred power could be introduced "if tech companies fail to live up to their new

responsibilities" and that a review would take place at least two years after the new regulatory regime was fully operational.

Recent comments by Johnson suggest the government may be considering strengthening the bill with additional criminal sanctions, which may include criminal liability for executives who fail to prevent wrongdoing. However, government sources have played down the remarks and denied plans to change the bill.

How this issue plays out when the bill is debated before Parliament remains to be seen.

3. Does the service provider need to be based in the UK?

As currently drafted, the bill will apply to service providers — regardless of where they are based — if they have links to the U.K.

A service provider will have links to the U.K. if:

- The service provider has a significant number — as yet undefined — of users in the U.K.;
- U.K. users form one of the service provider's target markets; or,
- The services are capable of being used in the U.K. and there are reasonable grounds to believe there is a material risk of significant harm to individuals in the U.K. arising from content on the service or that may be encountered in or via search results.

4. What is the purpose of the bill and what content will it regulate?

The bill will impose various duties on regulated service providers to take actions in respect of content that is either illegal content or content that is legal but harmful to adults or children. The bill contains some clarity on what these terms mean but much will be left to the discretion of the secretary of state, leaving scope for considerable uncertainty.

This uncertainty has been widely criticized, including for placing a disproportionate burden on service providers to fill in the blanks. The need for clarity will be especially important in circumstances where there are plans to introduce criminal liability for senior managers who fail to prevent breaches of the duties.

5. What is illegal content?

Illegal content has been defined as content consisting of certain words, images, speech or sounds whose use or dissemination is reasonably believed by a service provider to amount to:

- A terrorism offense;
- A child sexual exploitation and abuse offense;
- A priority illegal content offense, to be identified by the secretary of state; or

- Other offenses not within the previous three sets where the victim is an individual or individuals, to be determined by the secretary of state.

6. What is harmful content?

Content that is legal but harmful to adults or children has been defined as content which:

- The U.K. secretary of state designates as such; or
- The service provider believes is materially at risk of having a significant adverse impact on ordinary sensibilities of adults or children — e.g. pornography, violence, content regarding eating disorders, self-harm or suicide.

The U.K. government has indicated^[6] this category is likely to include user-generated fraud, including romance scams.

7. What are the duties on regulated service providers?

The bill imposes duties on regulated service providers, including:

- General duties that apply to all regulated service providers;
- Further duties that apply to regulated service providers whose services are likely to be accessed by children;
- A set of additional duties in relation to content that is harmful to adults and that apply to certain user-to-user service providers that fall under Category 1 — the threshold conditions for which will be set by the U.K. secretary of state; and,
- Various other duties that apply to certain search service providers and that fall under Category 2A and certain user-to-user service providers that fall under Category 2B — the threshold conditions for which will be set by the U.K. secretary of state.

We have produced a table that sets out in more detail the duties imposed on different types of service providers.

In setting the conditions for inclusion as a Category 1 user-to-user service provider, the secretary of state shall have regard to the risk of disseminating material that is harmful to adults and the number of users and functionalities of the services. This reflects the fact that providers of Category 1 services owe additional safety duties in respect of content which harmful to adults.

It is anticipated that Category 1 will be reserved for a small number of the most popular service providers that are considered high-risk and high reach.

In setting the conditions for service providers and user-to-user service providers to be included within Categories 2A and 2B, the secretary of state shall have regard to the prevalence of illegal content, content that is harmful to children or adults on the one hand, the number of users and functionalities of the services and any other factors the secretary of state deems relevant.

The regulator will then assess services against these thresholds and publish a register of all those service providers that meet the thresholds for inclusion in the various categories.

Duties Imposed on Different Types of Service Providers

Duties Imposed	All service providers	All service providers accessible by children	Category 1 service providers	Category 2A and 2B service providers
Risk assessment				
Conduct an illegal content risk assessment.				
Conduct an assessment to determine whether children have access to the service.				
Conduct a children's risk assessment.				
Conduct an adult's risk assessment.				
Conduct a risk assessment on the impact of policies and procedures on freedom of expression and privacy risk.				
Steps to mitigate and effectively manage risks				
Take proportionate steps to mitigate and effectively manage the risks identified in the illegal content risk assessment.				
Take proportionate steps to mitigate and effectively manage the risks of harm and impact of harm to children as identified in the children's risk assessment.				
Systems and processes				
Set up proportionate systems and processes to minimise the presence of priority illegal content and swiftly take down illegal content that the service provider is alerted to.				
Set up proportionate systems and processes to prevent children from accessing harmful content.				
Set up systems and process to protect democratic content.				
Set up systems and process to protect journalistic content.				
Terms of service				

Specify in the terms of service how individuals are to be protected from illegal content.				
Specify in the terms of service how children are prevented from encountering harmful content.				
Specify in the terms of service how content that is harmful to adults is dealt with.				
Specify in the terms of service the policies and processes that are designed to protect democratic content.				
Specify in the terms of service how journalistic content is identified and treated, as well as the policies and processes for handling complaints relating to journalistic content.				
Specify in the terms of service or in a publicly available statement, the positive steps taken in response to the freedom of expression and privacy risk assessment.				
Rights to freedom of expression and privacy				
Have regard to the importance of freedom of expression and privacy in deciding on and implementing safety policies and procedures.				
Reporting and complaints mechanism				
Operate a reporting service which allows users and affected persons to report content.				
Operate a transparent and easily accessible complaints procedure which allows for complaints regarding the operation of the systems and processes and provides for appropriate action to be take in relation to complaints.				
Create a dedicated and expedited complaints procedure available to a person who considers content taken down or restricted is journalistic content.				
Transparency				
Produce an annual transparency report.				

8. Who is the regulator, what are their enforcement powers and what guidance will they offer?

The existing U.K. regulator for communication services, Ofcom, is expected^[7] to be given additional powers to regulate the proposed regime.

Ofcom will be given the power to fine service providers failing in respect of a new duty of care up to £18 million (just under \$24.3 million) or 10% of annual global turnover, whichever is higher, and will have the power to block access to sites. The fines under this regime are potentially even higher than the significant fines that can be imposed under the General Data Protection Regulation, the maximum fines being the highest of €20 million (\$22.6 million) or 4% of a company's annual global turnover.

Ofcom is required to prepare a code of practice for service providers describing the recommended steps for compliance with the new duties set out in the proposed legislation.

Service providers whose revenue is equal to or exceeds a threshold figure to be set by Ofcom may need to pay an annual fee.

9. What is the status of proposed equivalent legislation in the U.S. and EU?

Reforms are also being proposed and/or debated in many other jurisdictions, including the U.S. and EU.

Over the last few terms, the U.S. Congress has introduced a number of bills aimed at content moderation reform. Many of these bills focus on specific content, such as child trafficking, terrorism, cyberbullying and illegal drugs. Others propose more wholesale reform of the current system of immunity that exists for online services that publish third-party content.

Many of these proposed wholesale reforms have a narrow applicability, generally based on revenue or market value, or appear to be directed specifically at large tech companies. Some states have also enacted content moderation laws.

Many of these laws have been challenged in court as preempted by federal law, but states are likely to continue to craft their own content regulation laws. This may — and indeed already has — result in a patchwork of laws throughout the United States.

Meanwhile, the European Commission proposed a "Digital Services Act Package"^[8] to the European Parliament and European Council on Dec.15, 2020, including introducing a legislative framework for moderating illegal or potentially harmful online content as well as various transparency requirements — for example, to notify posters if their content has been removed and to publish detailed reports of all content removed, liability of online intermediaries — e.g., hosting providers or online platforms for third-party content.

Unlike the proposed U.K. bill, the proposed Digital Services Act also includes mechanisms to warn users about fraudulent advertising.

10. What should companies do to prepare?

Although the bill is at a relatively advanced stage, it may yet be some time before it comes into force and it may be subject to change as it progresses through the U.K. Parliament. Even when it is enacted, much will hinge on the discretion of the U.K. government as to the meaning of harmful content and the

thresholds for the different categories of service providers.

Other proposed legislation in the U.S., the EU and other jurisdictions around the world is still some way off. Despite legislation being some way off, the matter has been a hot topic in the U.S. Congress. In addition, calls for action are not only coming from the authorities but also from the market, with social users staging blackouts and boycotts in response to failures to tackle online racial hate speech and some advertisers making spending contingent on the implementation of effective content moderation schemes.

It is therefore unsurprising that many service providers have not waited for legislation to be brought in and are taking proactive steps to self-regulate.

We may not know exactly what is required to comply with the new regimes nor is there yet a market standard for good behavior, but there are certain steps that companies could think about taking sooner rather than later, if they have not already, to protect their users and themselves.

These include:

- Conducting risk assessments to assess the extent to which their services may be accessible by children and the potential for harm to users from online content; and,
- Designing sophisticated systems and controls that can detect, remove and block harmful content, including escalating marginal calls, while also having regard to users' rights to freedom of expression and privacy.

It is becoming abundantly clear that the online landscape is changing and platforms are increasingly facing pressure to take more action on content moderation, striking the difficult balance between protecting the right to freedom of expression with tackling online harms.

Any laws that are made will need to offer workable solutions to organizations, both in navigating the cross-border issues faced by global platforms and in articulating clear and precise requirements, with actionable guidance on the path ahead.

Lucy Blake is special counsel and Karam Jardaneh is an associate at Jenner & Block LLP.

The opinions expressed are those of the author(s) and do not necessarily reflect the views of the firm, its clients, or Portfolio Media Inc., or any of its or their respective affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice.

[1] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf.

[2] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/973939/Online_Harms_White_Paper_V2.pdf.

[3] <https://www.bbc.com/news/uk-politics-58998635>.

[4] <https://www.lawcom.gov.uk/reforms-to-protect-victims-of-online-abuse-and-safeguard-freedom-of-expression-announced/>.

[5] <https://www.gov.uk/government/news/landmark-laws-to-keep-children-safe-stop-racial-hate-and-protect-democracy-online-published>.

[6] <https://www.gov.uk/government/news/landmark-laws-to-keep-children-safe-stop-racial-hate-and-protect-democracy-online-published>.

[7] <https://www.ofcom.org.uk/about-ofcom/latest/features-and-news/ofcom-to-regulate-harmful-content-online>.

[8] <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>.