SLAUGHTER AND MAY/

NOVEMBER 2023

ONLINE SAFETY ACT BECOMES LAW, BUT WITH PHASED ENTRY INTO FORCE

A version of this briefing first appeared in Privacy Laws & Business UK, Report Issue 130 (November 2023)

On 19 September, the text of the Online Safety Act was finally agreed by Parliament, some eighteen months after its first *reading*. It received Royal Assent on 26 October.

The Online Safety Act (OSA), which imposes a host of new duties on in-scope online services, will, according to the government, "*make the UK the safest place in the world to be online*". But the Act has been proven particularly divisive during its long passage. Criticised for being "*complex and incoherent*" and excessive in length, civil liberties groups have argued the Act is a threat to freedom of expression and privacy. The NSPCC, on the other hand, believes the OSA marks a "*new era for children's safety online*".

Unsurprisingly, such scrutiny has led to the Act evolving considerably over time. Most notably, the government shelved plans to criminalise "harmful" communications, and to require tech companies to address "legal but harmful" content for adults. Notwithstanding this, the Act stands at over three hundred pages, and significant questions remain as to how workable it will be in practice.

In this article, we provide an outline of the new regulatory regime, including certain key duties it establishes and Ofcom's enforcement powers. We also discuss the relevance of privacy to the Act, and briefly set out the next steps for its phased entry into force.

Who is within the scope of the new regime?

The two key categories of internet services within the ambit of the OSA are **user-to-user services** and **search services**, although the Act does also regulate certain online pornographic content.

By "*user-to-user services*", the Act means internet services which contain the functionality for at least one user to encounter content uploaded to, shared on, or generated on, the service by another user. Social media sites, online gaming platforms, and video-sharing sites will therefore be among the services caught. "*Search services*" refers to services which contain a "*search engine*", that is, the functionality to search multiple websites or databases.

However, services will only be in-scope if they have "*links with the United Kingdom*". This means the service either: (i) has a significant number of UK users, or they form a target market for it; or (ii) can be accessed in the UK, and there are "*reasonable grounds*" to believe there is "*a material risk of significant harm*" to users from its content. The OSA therefore (as with GDPR) has extraterritorial scope; but importantly, the Act's provisions make clear that it is internally focused, being concerned only with the impact the service has on UK users, and on the design, operation or use of the service in the UK.

Furthermore, certain services (or parts of services) are specifically exempted from the OSA's scope. This includes SMS, MMS and e-mail services, services that are used as internal business tools, and "*limited functionality services*", being those whose only user-to-user interaction is via comments on a service provider's content. Nevertheless, the government has estimated that 25,000 organisations remain in-scope in the UK, with Ofcom putting the figure at over 100,000 when including overseas providers.

What are the new duties on service providers?

Under the current regime (set out in the E-Commerce Regulations 2002), the liability of service providers for unlawful content is reactive, that is, they must "*act expeditiously*" to take such content down once aware of it.

By contrast, the duties of care imposed by the OSA are broader and largely proactive in nature. Furthermore, while there is a 'sliding scale' of duties depending on the characteristics and scale of the provider, and whether

the service is accessible by children, all in-scope services will face a significant compliance burden.

The general duties in the OSA include:

- **Illegal content risk assessment**. Providers must conduct a "*suitable*" and "*sufficient*" assessment that, among other things, assesses the risk of users encountering illegal content on a service, and the risk and severity of harm they may face from such content. This assessment must be kept up-to-date, and a further assessment will be required each time a service makes any "*significant change*" to its design or operation. This duty is a foundational one, as it is this risk assessment that closely informs the measures adopted by a provider to tackle illegal content.

- **Safety duties concerning illegal content**. Providers must use "*proportionate*" measures, processes and systems to, among other things: (i) in the case of user-to-user services, prevent users encountering particular categories of "*priority*" illegal content (such as child abuse content), and "*swiftly*" take down any illegal content once on notice; (ii) in the case of search services, minimise the risk of individuals encountering illegal content; and (iii) for both types of service, "*effectively*" mitigate the risks of harm to individuals.

- **Content reporting and complaints**. Users must be able to easily report illegal content, and providers must operate complaints procedures that are easy to access, easy to use (including by children), transparent, and which provide for appropriate action to be taken. Users must be able to complain about, among other things, the removal or de-prioritisation of their content and actions taken against them, as well as non-compliance by a provider with their duties.

'Category 1' services (which are yet to be designated, but which will include the biggest social media companies) will be subject to further duties. For example, they will be required to "*empower*" adult users to filter the content and users they encounter, and must use proportionate systems to protect "*content of democratic importance*" as well as "*journalistic content*" (each defined in the Act).

The Act will also require certain providers to use reasonable measures to prevent individuals from encountering fraudulent adverts, and introduces (or bolsters) criminal offences relating to cyber-flashing, revenge porn, encouragement of self-harm, and "*threatening*" or "*false*" communications.

Are there additional duties regarding children?

The protection of child users is a key focus of the legislation. As such, there are additional duties for those services "*likely to be accessed by children*".

These include:

- conducting a specific "*children's risk assessment*", which includes considering the risk of harm to children of different age groups presented by "*content harmful to children*";

- using proportionate systems to prevent any child encountering "*primary priority*" content harmful to children (which includes suicide, self-harm or eating disorder content);

- protecting children in age groups judged to be at risk at harm from other harmful content; and

- generally mitigating the impact of harm on children (across all age groups).

Significantly, user-to-user and search services that have or are likely to attract significant number of child users cannot avoid these duties unless they use age verification or estimation techniques such that children cannot normally access the service. Services hosting regulated pornographic content, or those whose terms of service do not prohibit all "*primary priority*" content harmful to children, will also be mandated to use "*highly effective*" age verification to manage children's access to content.

How stringent are the duties on protecting users from certain content?

While providers will, as noted, be obliged to swiftly take down certain content upon notice, it is important to note that providers will not be expected – by virtue of the OSA's proactive duties – to protect every user or child from every piece of illegal content or "*content that is harmful to children*". This is because the duties in the OSA concerning illegal and 'harmful' content are qualified by proportionality. In ascertaining what is proportionate, the statute sets out two relevant factors: the size and capacity of the provider, and the results of its most recent risk assessment(s). However, the regime is focused at a structural level across the systems, measures and processes of a service. The duties will therefore impact organisation-wide, from the way algorithms are designed, to a provider's approach to content moderation and internal staff policies, and are essentially asking that a provider do what is appropriate and reasonable in the circumstances.

How will the OSA be enforced?

The OSA introduces a considerable enforcement and liability regime and Ofcom is the appointed regulator.

Most significantly, Ofcom can impose substantial financial penalties (exceeding those under the GDPR), with fines up to the greater of £18 million, or 10% of global annual revenue.

Ofcom's powers also include issuing "*service restriction orders*" and "*access restriction orders*" to essentially shut down a non-compliant service, and extensive powers to inspect a provider's premises (including potentially without warrant), demand information, conduct investigations and interviews, and audit providers.

Furthermore, the OSA includes the possibility of personal criminal liability for 'officers' (i.e. directors, managers), for example for a failure to comply with children's safety duties. This applies when the offence is committed by an entity "*with the consent or connivance*" of that officer, or where it is "*attributable to any neglect*" of that officer.

What is the relevance of privacy to the OSA?

The ICO and Ofcom have jointly recognised that there is a tension between protecting individual users from online harms on the one hand, and safeguarding privacy on the other. They have noted that tools used to ensure greater online safety could also involve more monitoring of user activity and collection of personal data to identify harmful behaviour or those who may be vulnerable. The OSA also envisages additional processing of personal data: Category 1 services are required to offer users the ability to verify their identity, and it is likely that in-scope providers will conduct age verification checks.

Despite this potential tension, the regulators have made it clear that they will expect service providers that are in scope of the OSA to meet both their online safety and data protection responsibilities. They confirm that "*there can be no space for services to argue that they could not comply with new online safety requirements, because of data protection rules, or vice versa.*"

The Act addresses the issue of privacy in a general way in relation to the vast majority of in-scope providers. It does so by imposing on providers a 'cross-cutting' duty to "*have particular regard to the importance of protecting users from a breach of any statutory provision or rule of law concerning privacy*", including those relating to the processing of personal data, when implementing "*safety measures and policies*" to satisfy certain identified duties (which includes the safety duties, content reporting and complaints, and child safety duties).

It may not, at first glance, be completely clear to a provider how to fulfil such a duty, or if having "*particular regard*" to privacy means anything other than continued compliance with relevant data protection laws (for example, conducting a DPIA if processing biometric data to estimate a user's age or using innovative technologies (such as AI) to manage its OSA obligations).

However, there is an important clarification in the Act that a service provider will be compliant with this duty by using such of the "*relevant recommended measures*" (being those set out in certain Ofcom codes of practice) "*as incorporate safeguards to protect the privacy of users*", to the extent they are relevant to the provider and service in question. As such, in-scope providers can take some comfort that detailed guidance will follow, and from the fact that the ICO and Ofcom have publicly committed to work together to "*provide a clear and coherent regulatory landscape*".

For the limited number of Category 1 services that will exist, there are more specific and onerous obligations relating to data privacy. These include an obligation to conduct, and make public, an assessment of the impact their safety measures will have on privacy (as well as freedom of expression), and keep it up to date.

Notwithstanding the inclusion of such duties, the Act has been criticised on the grounds that its robust safety duties outweigh privacy considerations. For example, human rights organisation Article 19 have argued that the duty to "*have regard*" to privacy offers little meaningful protection, and that reliance on Ofcom codes "*removes any incentive for a company to duly consider the impact of its safety measures*" and may reduce it "*to a box-ticking exercise.*"

There has also been widespread public scrutiny of the power the OSA grants Ofcom to require in-scope providers to use or source technology to scan private messages for child abuse content. This prompted a number of tech companies to threaten withdrawal from the UK market, with companies such as WhatsApp and Signal arguing that scanning technology does not, and could not, exist without removing end-to-end encryption for all users and undermining their privacy. In light of the backlash, the government has since sought to clarify that it has no intention of "*weaken[ing] the encryption technology used by platforms*". It also explained that any such technology would need to meet standards set out in the Act, and if no such technology is developed, Ofcom clearly cannot demand its use. While some have viewed this as a concession that the power will not in practice be

exercised, the relevant parts of the legislation remain unchanged.

What will happen next?

Ofcom, the appointed regulator, expects its powers to commence two months after the OSA becomes law.

It will then begin carrying out a series of phased consultations relating to draft Codes of Practice and guidance. **Phase one** will focus on illegal harms, and will be published shortly after Ofcom's powers commence. **Phase two** will focus on child protection duties, and is expected to be published approximately six months after Ofcom receives its powers. **Phase three** will focus on the particular requirements that fall upon Category 1 and Category 2 services, with Ofcom expecting to submit its advice to the government on how it should categorise services within six months from Royal Assent.

Ofcom's expectation is that the safety duties concerning illegal content will enter into force, and its first Codes of Practice will be issued, approximately one year from its first consultation commencing, with companies expected to conduct their risk assessments in the months preceding issuance.

The government has also signalled that it may address any gaps identified in the OSA through changes to the Data Protection and Digital Information Bill.

In terms of the ICO, the statute requires that Ofcom consults the ICO on each Code of Practice it prepares and before issuing guidance. Furthermore, the ICO will itself issue guidance on the data protection expectations for online services using safety technologies. It will also monitor online harms more generally, taking into account its existing guidance, such as the Children's Code.

CONTACT

ROB SUMROY
PARTNER
T: +44 (0)20 7090 4032
E: rob.sumroy@slaughterandmay.com

JACK HIGGINS
ASSOCIATE
T: +44 (0)20 7090 5111
E: jack.higgins@slaughterandmay.com