ONLINE SAFETY ACT 2023
A REVOLUTION IN REGULATION

Laura De Boel, Tom Evans and Hattie Watson of Wilson Sonsini Goodrich & Rosati highlight some of the key features of the Online Safety Act 2023, together with the challenges that companies may face when looking to comply with its requirements.

The Online Safety Act 2023 (2023 Act) was enacted in October 2023, concluding its slow and, at times, dramatic passage through Parliament. The 2023 Act, which will impose extensive new regulatory requirements on an estimated 100,000 companies that offer user-to-user services or search services, has been heralded by the government as a development that will "make the UK the safest place in the world to be online". On the other side of the debate, however, the 2023 Act has been decried as an affront to freedom of speech that will burden companies with disproportionate obligations.

Whichever way it is approached, the 2023 Act is one of the most significant pieces of legislation affecting user-facing services to be enacted in a generation.

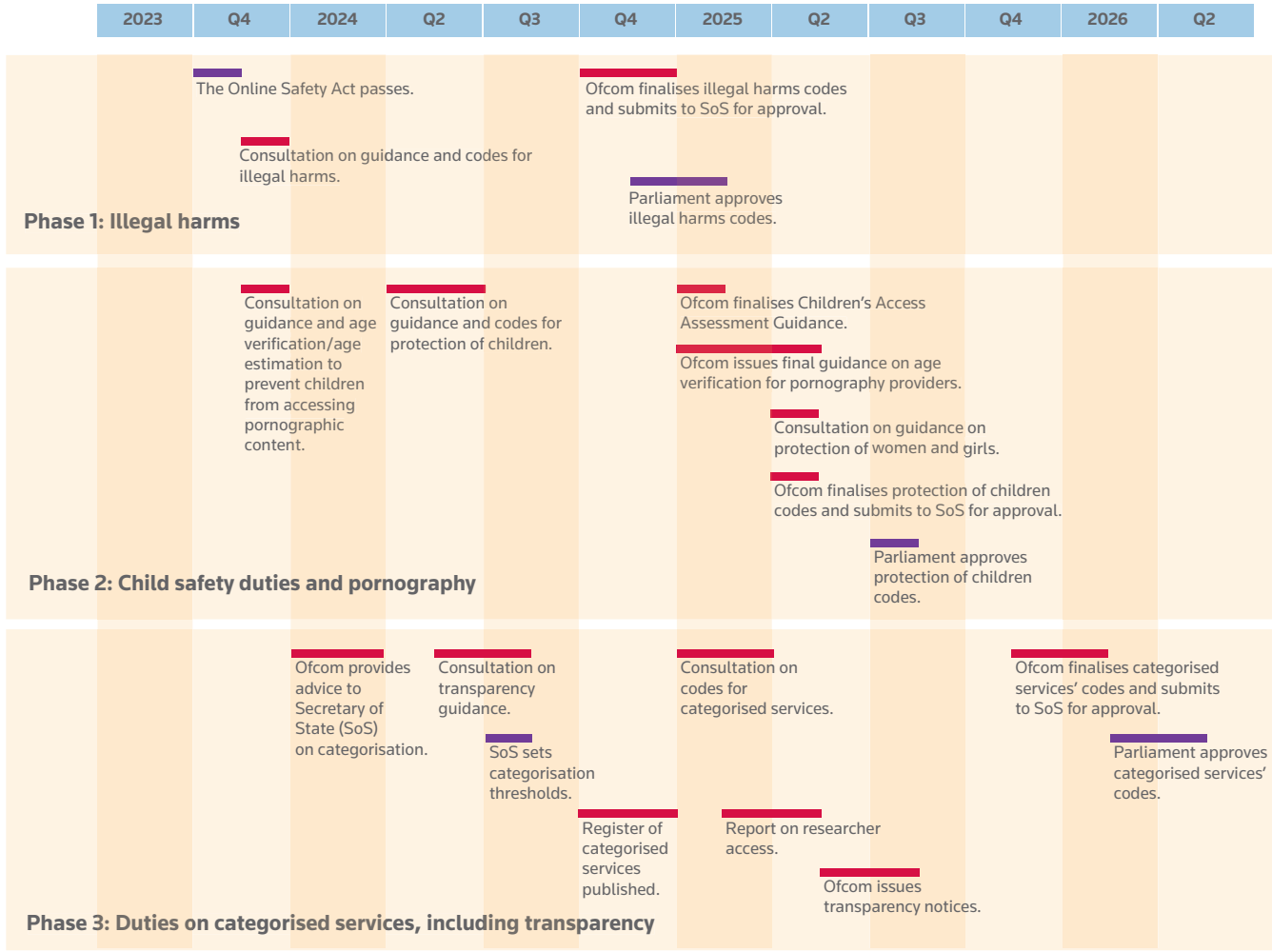
This article highlights some of the key features of the 2023 Act, together with the challenges that companies may face when looking to comply with its requirements. It also highlights some of the key differences between the 2023 Act and the EU's flagship Digital Services Act (DSA), which entered into force in 2022, and will become applicable to most in-scope services in early 2024 (*see feature article "Regulating digital services in the EU: a paradigm-shifting legislative framework", www.practicallaw.com/w-030-6172*).

OBJECTIVE OF THE 2023 ACT

The 2023 Act's stated aim is to make regulated internet services safer for individuals in the UK, in particular by requiring companies to assess and manage safety risks arising from content and conduct on their services. It responds to concerns that were identified by the government during a series of consultations on online harms. These consultations found that 61% of adults and 79% of 12 to 15-year-olds have had at least one potentially harmful experience online within the preceding year (*www.ofcom.org.uk/__data/assets/pdf_file/0028/149068/online-harms-chart-pack.pdf*). Following the consultations, the government concluded that legislative action was required to address evidence of both illegal and harmful content and activity taking place online (*see Briefing "Online harms: shaping a safer digital world", www.practicallaw.com/w-025-6448*).

The 2023 Act's focus on online safety is reflected in the types of content that it looks to regulate; it is primarily concerned with content and conduct that amounts to a

Phased implementation: government steps



Source: Ofcom's approach to implementing the Online Safety Act, 26 October 2023

criminal offence. In relation to child users, it seeks to regulate a wider category of "harmful" content and conduct.

This focus on content that amounts to a breach of criminal law is one important feature that distinguishes the 2023 Act from the DSA. Whereas the DSA looks to tackle all forms of "illegal" and infringing content, the 2023 Act does not aim to regulate content that amounts to an infringement of civil law, such as defamation or the infringement of intellectual property rights. The existing statutory notice and takedown regime relating to those types of content will not be altered by the coming into force of the 2023 Act.

While the 2023 Act's overall objective is to improve the safety of online services, the online safety regulator, Ofcom, has made it clear that it does not expect harmful and illegal content to be entirely eradicated. This is reflected in the risk-based nature of the obligations that are imposed on service providers by the 2023 Act.

THE ROLE OF OFCOM

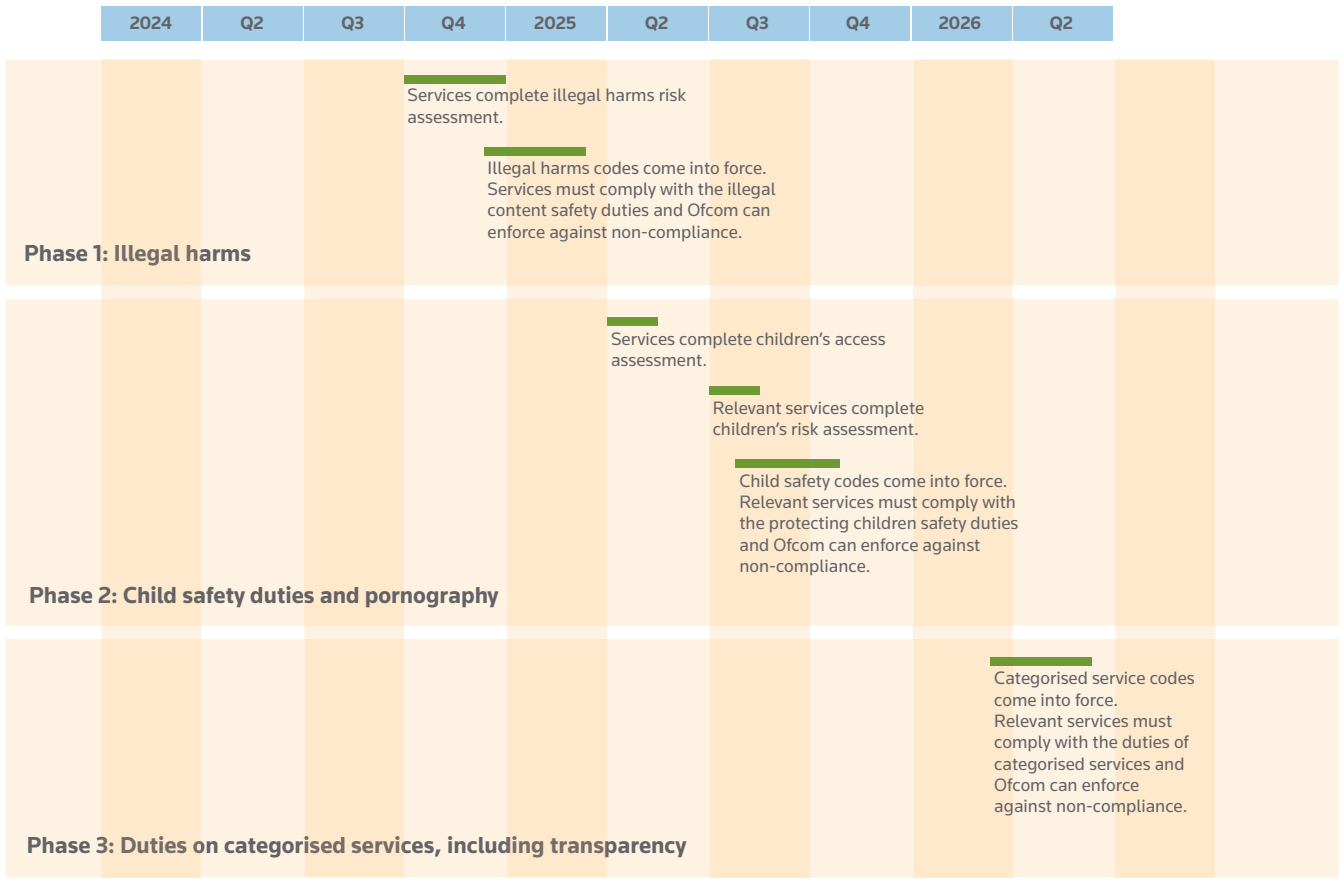
Ofcom has a central role in both bringing the duties on companies under the 2023 Act into effect, and in subsequently supervising compliance.

Bringing the duties into force

Ofcom is responsible for publishing guidance and codes of practice, and providing direction on the measures that companies are required to take to comply with their obligations under the 2023 Act. These documents will first be issued for public consultation. When finalised following consultation, the codes of practice will be laid before Parliament before coming into force, and guidance notes will be published in official form by Ofcom. The first public consultation, titled "Protecting people from illegal harms online", was published on 9 November 2023 (the November consultation) (*www.ofcom.org.uk/__data/assets/pdf_file/0020/271145/volume-1-illegal-harms-consultation.pdf*).

The duties under the 2023 Act will come into force as the first pieces of guidance and codes of practice are finalised. Ofcom has published a detailed Roadmap to Regulation, which sets out its expectations as to when the guidance and codes will be published. However, there is scope for potential slippage

Phased implementation: milestones for service providers



Source: Ofcom's approach to implementing the Online Safety Act, 26 October 2023

due to political events, such as the general election in 2024 or early 2025 (*see boxes "Phased implementation: government steps" and "Phased implementation: milestones for service providers"*) (*www.ofcom.org.uk/__data/assets/pdf_file/0017/270215/10-23-approach-os-implementation.pdf*).

Enforcement

Ofcom has broad powers to investigate and ultimately impose penalties on providers that it considers have not met their obligations under the 2023 Act. In addition, where Ofcom has grounds to do so, it can apply to the courts to obtain orders to restrict the provision of services in the UK, or to restrict access to the infrastructure that allows the services access to the UK market, such as an internet service provider or app store. In the case of the largest platforms, some of which will already have been subject to regulation by Ofcom under the video-sharing platform regime, Ofcom has stated that it will follow a model of continuous regulatory supervision, involving regular engagement with providers.

Monetary penalties under the 2023 Act can amount to up to £18 million, or up to 10% of annual worldwide revenue, whichever is greater. This exceeds the maximum penalty for infringements of the DSA, which is set at 6% of total worldwide annual turnover.

Perhaps most significantly, the 2023 Act provides that senior managers can be held criminally liable for certain infringements of the legislation. For instance, where Ofcom issues an information notice to a regulated business, it may require that entity to name a senior manager who is in a position to ensure compliance with the requirements of that notice. Once named, the individual may be criminally liable if the business fails to comply with the notice. Where Ofcom exercises its power to conduct interviews under the 2023 Act, an offence may also be committed if an individual fails to attend, or provides information that is false in a material respect.

SERVICES CAUGHT BY THE 2023 ACT

The main part of the 2023 Act (Part 3) regulates two categories of online services:

- User-to-user services.
- Search services.

For the 2023 Act to apply, these services must have "links with the UK", either because they have a significant number of users in the jurisdiction, target the jurisdiction as a market, or otherwise present a material risk of significant harm to individuals in the jurisdiction.

User-to-user services

An online service will amount to a user-to-user service if it allows users to generate, upload or share content that may be encountered by

another user, provided that an option to share the content exists. This will include the largest online platforms and social media companies, but also smaller providers of interactive online services or providers of messaging services. Online businesses offering dating apps, online gaming or marketplaces are also likely to be caught.

Search services

A search service is any platform that has some function that allows an individual to search multiple websites; for example, an internet search service, a travel comparison website or a search function in an academic article database.

Unlike the DSA, the 2023 Act does not contain explicit exceptions for small and micro enterprises. While a limited number of exceptions are available for specific types of services, such as certain services that are provided in the context of education and healthcare, Ofcom has noted that "all in-scope user-to-user and search services, large or small, will need to take action to comply with the new duties".

DUTIES ON COMPANIES

The 2023 Act seeks to achieve its objective by imposing a range of duties on the companies that fall within its scope. The guidance and codes of practice that are to be published by Ofcom will provide companies with an indication of the measures that it expects to be taken to address compliance with specific duties. Codes of practice will be disclosable in evidence in legal proceedings, presumably to allow the courts to assess a service provider's conduct against the standards set by Ofcom.

Companies will be treated as complying with a duty if they take or use measures described in a code of practice. Companies that opt to implement alternative measures will bear the burden of proving that those measures are sufficient to meet their duties under the 2023 Act.

Phase 1 illegal harms

Under the 2023 Act, "illegal content" is defined as content that amounts to an offence and "priority illegal content" means offences that are most harmful, including terrorism content, and child sexual exploitation and abuse material. The regulatory framework set out in the 2023 Act is designed to require companies to assess and manage the risks to which users are exposed, in particular from illegal content and activity, and content and activity harmful to children.

The legislation does this by requiring in-scope service providers to carry out risk assessments that must consider the likelihood of specific types of illegal content being available on their service, along with the risk that the service may itself be used to commit certain offences. An earlier aim for the 2023 Act to protect users more generally against harmful content was dropped during the legislative process.

Having identified these risks, service providers are required to implement mitigation measures that are tailored to the risks identified. Illegal content duties will be among the first obligations under the 2023 Act that providers will need to comply with. Ofcom published draft guidance and codes of practice as to how platforms may comply in the November consultation.

Risk assessment. Illegal content risk assessments will be expected to meet the complexity of the services to which they relate. In the course of carrying out an illegal content risk assessment, companies must assess the risk that users will encounter priority illegal content, as well as other illegal content on their services. In the November consultation, Ofcom has proposed a four-step process for illegal content risk assessments, under which companies will:

- Understand the harms present on their services by reference to Ofcom's published risk profiles.
- Assess the risk of each type of harm and assign a risk level to each one.
- Decide on appropriate measures to reduce the risk of harm to individuals and implement these measures.
- Report, review and update risk assessments, and monitor the effectiveness of the mitigation measures that were implemented.

Risk mitigation measures. The "appropriate measures" that companies should implement to reduce the risk of harm to individuals using their service will vary. However, the 2023 Act provides a list of example measures (*see box "Example mitigation measures"*).

The November consultation provides a glimpse of Ofcom's expectations when it comes to these measures. In particular, it indicates that services of all sizes should appoint a named person to be accountable to the most senior governance body of the relevant company (which in most cases will be the board) for compliance with illegal content safety duties, and reporting and complaints duties. Senior members of staff that make decisions related to the management of risk should have written statements of responsibility and most companies will be

Example mitigation measures

The Online Safety Act 2023 sets out examples of mitigation measures that service providers can take:

- Regulatory compliance and risk management arrangements.
- The design of functionalities, algorithms and other features.
- Policies on terms of use.
- Policies on user access to the service or to particular content present on the service, including blocking users from accessing the service or particular content.
- Content moderation, including taking down content.
- Functionalities that allow users to control the content they encounter.
- User support measures.
- Staff policies and practices.

expected to maintain their own tailored internal content moderation policies that are assessed by reference to performance targets. Services that are at specific risk of child sexual exploitation and abuse offences may be expected to take additional measures to reduce the risk to individual users, including implementing specific technologies to detect images.

This wide-ranging requirement for all companies to conduct an illegal content risk assessment distinguishes the 2023 Act from the DSA in the EU. Under the DSA, only the largest platforms (known as "very large online platforms" or "very large online search engines") are required to conduct a risk assessment, which must focus on identifying systemic risks.

Content relevant to duties relating to children

Under the Online Safety Act 2023, the first tier of content is "primary priority content that is harmful to children", which refers to content that is considered to be most harmful, including content that is pornographic or that encourages, promotes or provides instructions for suicide, an act of deliberate self-injury or an eating disorder and associated behaviours.

"Priority content that is harmful to children" is the second tier and includes, for example, content that is abusive and targets characteristics that are protected or incites hatred against people with such characteristics, such as race, religion and sex. In addition, it covers content that is bullying or that encourages, promotes or provides instructions for a challenge or stunt that is highly likely to cause serious injury.

Non-designated content that is harmful to children refers to any other content that presents a material risk of significant harm to children in the UK. Ofcom will produce guidance, with examples, of how platforms can assess how to label content.

Phase 2 children

A major focus of the 2023 Act is on protecting children online, with several amendments having been made during the legislative process that are intended to increase the protections afforded to individuals under the age of 18. Companies that fall within the scope of the 2023 Act will be required to consider the use, or potential use, of their services by children as part of a multi-step process, even if they are not directed or intended for them.

Access assessment. All service providers will be required to carry out a children's access assessment to establish whether their service is likely to be accessed by children in the UK. As part of this assessment, providers will need to identify whether a significant proportion of their UK service users are children, or whether the service is likely to attract a significant number of child users in the UK. Providers will only be able to conclude that their services, or parts of their services, are not likely to be accessed by children if they implement some form of age assurance, such as age verification or age estimation, in order to prevent such access (*see "Age estimation and verification" below*).

The children's access assessment is an ongoing duty, so providers, whether their service is likely to be accessed by children or not, will need to renew the assessment at least annually or in light of changing circumstances; for example, before making significant changes to the service's design or operation. Ofcom will publish draft guidance for consultation on how providers can conduct this assessment in 2024 and it is likely that providers should be prepared to conduct their first assessment in early 2025.

Children's risk assessment. If a service, or part of a service, is likely to be accessed by children, providers will need to conduct a children's risk assessment. This separate and additional assessment is wide-ranging and focuses on the online harms that children of different age groups could face on the service. In particular, the assessment should identify the risk posed by encountering "content that is harmful to children". This umbrella term is split into three subcategories according to the level of harm that it could cause (*see box "Content relevant to duties relating to children"*).

The risk assessment will also need to account for the user base of the service, including the number of children in different age groups, the impact of algorithms, the design of the service and how the service could be used.

Children's mitigations. Having identified the risks that their service presents, providers must implement measures to protect children. This will involve taking measures to:

- Protect children of all age groups from encountering primary priority content that is harmful to children.
- Mitigate and manage the risks of harm to children that were identified in the most recent risk assessment.
- Mitigate the impact of harm to children presented by any content that is harmful to children on the service, including by protecting children in age groups that are judged to be at risk of harm from encountering the content.

Providers will be expected to tailor policies and measures to account for how different age groups may be affected by particular content. In particular, providers must recognise that some age groups may be at greater risk of harm from particular content than others.

Providers have a range of tools at their disposal to comply with these safety duties, including content moderation, regulatory risk management and functionality that allows children to control the content that they view. Applying age assurance techniques to age-gate particular content is highlighted as one option. In fact, this is required where the providers' terms of service do not explicitly prohibit all users from uploading or generating primary priority content that is harmful to children on the service. Providers will be required to add details to their terms of service to explain to users how they comply with the requirements, including the use of any proactive technology.

Companies should note that the DSA also requires online platforms to implement measures to ensure the privacy, safety and security of minors using their service. While the DSA itself is less prescriptive about how this objective should be achieved, it is anticipated that regulatory guidance will be adopted in the medium term. By way of example, there is no requirement under the DSA to carry out a separate children's risk assessment.

Phase 3 categorised services

Under the 2023 Act's framework, services that meet certain criteria relating to their number of users, functionalities and other characteristics can be designated as category 1 services, with other services to be designated as category 2A or category 2B services. Companies falling into these categories are required to comply with the most extensive obligations under the 2023 Act. These obligations include:

- Enhancing transparency. Providers of category 1 and category 2A services will be required to publish information about the findings of their most recent illegal content risk assessment and, where applicable, the children's risk assessment, including the nature of the risks, their severity and the potential harm to individuals. These providers also have obligations to supply a record of their risk assessment to Ofcom in full.

- Empowering adult users. Providers of category 1 services must ensure that their risk assessment examines the range of content that adult users of their services may encounter. They are subject to an additional duty to include features that allow users to increase their control over the types of content that they encounter, including, for example, content that is abusive on the grounds of race, religion or gender, or content that encourages or promotes suicide or self-injury.

- Offering identity verification. Users of category 1 services must be offered an option to verify their identity, if this is not already required for access to the service (*see "Age estimation and verification" below*). Transparent information about how verification works must be set out in the terms of service.

- Protecting certain categories of content. When considering whether to take action against content or against a user who generates and shares content, category 1 service providers are required to:

  - take into account the importance of freedom of expression in relation to content that is of democratic importance;

  - provide news publishers with advance notice and an opportunity to make representations; and

  - respect the importance of free expression of journalistic content.

Age estimation and verification

Technologies and processes that allow companies to verify or estimate the age of users will take on particular significance under the 2023 Act. Services that host pornographic content will be required to use age verification or estimation in order to prevent child access. All other services should consider the implementation of these technologies and processes in order to support their child access assessment, and to subsequently assess and mitigate the risks of harm to children. Indeed, in the future, Ofcom may recommend their use in this way.

While the precise requirements that these technologies must meet in order to be considered effective are not set out in the 2023 Act, the legislation makes clear that processes that rely on a user self-declaring their age will not be regarded as sufficient to comply with the 2023 Act's requirements. Ofcom will provide guidance on age assurance in a code of practice; draft guidance on preventing children from being able to access pornographic content is due to be released for consultation in December 2023.

Communications offences

The 2023 Act also introduces newly formulated communications offences:

- Sending false communications, where the person sending the message has an intent to cause non-trivial psychological or physical harm to the likely audience.

- Sending threatening communications of death or serious harm, with an intention to cause fear that the threat will be carried out, or recklessness as to that effect.

- Sending or showing flashing images electronically, to counter epilepsy trolls.

- Encouraging or assisting serious self-harm.

LOOKING AHEAD

Through enacting the 2023 Act, Parliament has made clear its broad expectations as to

Related information

This article is at practicallaw.com/w-041-4411

Other links from uk.practicallaw.com/

Topics

E-commerce	topic/2-103-1274
Internet	topic/8-383-8686
Social media	topic/0-525-4280
Telecoms	topic/7-205-8953

Practice notes

Digital Services Act (EU): overview	w-038-4350
Online platforms: dealings with consumers and business users	w-021-0867
Online Safety Act 2023	w-030-3139
Offences using an electronic communications network or service	w-019-3831
Social media compliance	w-020-2218
Social media: offences and civil causes of action	3-616-4951
Video-sharing platforms	w-024-1729

Previous articles

Digital markets regulation: comparing the new EU and UK regimes (2023)	w-040-0659
EU regulatory data framework: a new generation (2022)	w-036-5428
Regulating digital services in the EU: a paradigm-shifting legislative framework (2021)	w-030-6172
Online content: managing the growing youth market risk (2011)	6-505-8573

For subscription enquiries to Practical Law web materials please call +44 0345 600 9355

the steps that companies should take when offering services to users in the UK. However, much of the detail is yet to be worked out, and over the coming months companies should seize the opportunity to engage with Ofcom and its consultations, in order to help to shape the compliance requirements that they will be assessed against in the coming years.

In the meantime, companies can prepare for the phased implementation of the 2023 Act by gathering data on their user base, in particular the age groups of users, and reviewing current risk management activities. Having well-drafted and implemented terms of service will be important in complying with the 2023 Act and companies can start to consider the likely updates that will be required when the time comes.

Laura De Boel is a partner, Tom Evans is a senior associate, and Hattie Watson is a trainee solicitor, at Wilson Sonsini Goodrich & Rosati.