



A GUIDE TO THE ONLINE SAFETY ACT AND ITS IMPLEMENTATION

March 2025

This note provides an overview of the Online Safety Act, with links to further detailed information on [our website](#), and provides an update on its current implementation status, gaps and next steps. Please do contact us for further information.

1. The Online Safety Act: where we are now

The [Online Safety Act](#) received Royal Assent on 26 October 2023 after a long and quite tortuous [six-year passage](#) from policy development to the statute book. A short guide to the Act and what it does is [here](#) and we summarise the main duties below.

As the designated regulator, Ofcom has received its powers under the Act but the statutory consultation requirements and consequent long timescales mean that the main duties on regulated companies are only now beginning to come into force. The regime will only be fully in force once all the supporting codes of practice are published and the relevant statutory instruments have been laid.

As of February 2025, Ofcom has completed two major consultations on illegal harms and protection of children respectively. It issued its [illegal harms code](#) in December 2024 and this, as required by the Act, was laid before Parliament. It published its [child access assessment guidance](#) for Part 3 services and [age assurance guidance for part 5](#) (pornography service) providers in January 2025 and its [child protection code](#) is due to follow in spring 2025. Consultations on its draft transparency guidance and information-gathering powers also took place during 2024. Consultation on VAWG Guidance is to be launched on 25 February 2025. A full implementation timeline is provided at section four below and on our website: [here](#).

In addition to and separate from the regulatory provisions, the Act also introduced some new criminal offences or amended existing offences. These are found in [Part 10](#).

2. The Online Safety Act: the duties on regulated services

a) Illegal content

Illegal content duties are triggered when there is illegal content on a service. Illegal content comprises two categories: non-designated content and priority content. Non-designated content is anything that satisfies a general definition in s 59(5) OSA. Priority content is found in Schedules 5-7: these include terrorism offences ([Schedule 5](#)), child sexual exploitation and abuse offences ([Schedule 6](#)) and priority offences ([Schedule 7](#)), including eg threats to kill, public order offences, fraud, stalking and harassment; human trafficking and sexual exploitation.

In relation to all these offences, user-to-user services are required to:

- Effectively mitigate and manage risk of harm to individuals from illegal content
- Operate a system to allow swift take down of priority illegal content;
- Have a complaints process.

For priority offences, there are additional obligations on user to user service providers to:

- Mitigate the risk of the use of service for the commission of an offence
- Minimise the time illegal content content is present
- Prevent users from encountering that content

The duties vary in some respects in relation to user-to-user services and search services. While search services are under a duty to mitigate risk and to minimise visibility of priority content, there are no take down duties.

For all services, there are also a number of cross-cutting duties including eg record keeping and review. Guidance on content and activity which disproportionately affects women and girls should also inform services' approach to implementation of their illegal content and children's safety duties; its [draft guidance](#) is currently out for consultation.

For more detail on the illegal content duties and how they operate, please see our explainers on the [user-to-user duties](#) and the [search duties](#), as well as our [guide to the schedules of offences](#).

Implementation status: Ofcom consulted on the illegal content codes and other guidance between November 2023-February 2024 and published their statement, plus final risk assessment guidance and codes for Parliamentary approval, in December 2024.

In force?: the illegal harms risk assessment obligations came into force in December 2024. Regulated services have to finalise their first risk assessments by 16 March 2025, which is when the codes of practice are expected to come into force, following Parliamentary approval, and services will be required to comply with them. Ofcom [announced on 3 March](#) it was starting its enforcement programme, and had issued "formal information requests to providers of a number of services with a deadline of 31 March by which to submit their records of their illegal harms risk assessments to us."

b) [Protection of children](#)

All part 3 services (user-to-user and search) must undertake a Children's Access Assessment (CAA). If a service is likely to be accessed by children, it must complete a "suitable and sufficient" risk assessment and implement safety measures (including age assurance).

The OSA's children's safety duties apply to three categories of content: "primary priority"; "priority" and "non-designated".

- [Primary priority](#) includes pornography; content promoting or providing instructions for suicide, self harm or for an eating disorder. **The duty here is to prevent (or for search to minimise the risk of) children from encountering**
- [Priority](#) includes abuse/hatred on the basis of protected characteristics; Bullying content; violent content; instructions for dangerous stunts. **The duty here on user to user is to protect children at risk of harm from encountering; search and user to user must mitigate risks of harm**
- Non-designated is any content "of a kind which presents a material risk of significant harm to an appreciable number of children". **The duties are the same as for priority content.**

For more detail on the child safety duties and how they operate, see [our explainer](#).

Implementation status: Ofcom consulted on the details between March-May 2024. Children's access assessment guidance and age assurance guidance were published in January 2025. The final children's risk assessment guidance and protection of children codes are due in April.

In force?: all services must carry out a children's access assessment by 16 April 2025. When the risk assessment guidance is published in April, services likely to be accessed by children will have three months to carry out their risk assessment before the children's codes of practice are likely to come into force (July 2025).

c) [Pornography services](#)

Separate duties are placed on [part 5 services](#) (commercial porn providers): they must implement "highly effective" age verification or assurance to prevent children encountering pornographic content on their services. They also have to comply with record-keeping duties.

Implementation status: Ofcom consulted on the age assurance proposals between December-February 2024 and the final guidance for part 5 services was published in January 2025.

In force?: yes, the part 5 duties came into force on 17 January 2025. Services must implement age verification or age estimation on their services by July 2025.

d) [Additional duties](#)

Services will be categorised (according to regulations laid by DSIT but - as at 20/2 - not yet approved¹) as follows:

- Category 1: large user-to-user services - with a recommender system and a threshold UK monthly users of either 34m+ or 7m+ if the service has a user forwarding / sharing content functionality
- Category 2A: large search services - 7m+ monthly UK users
- Category 2B: smaller U2U services - 3m+ users with a direct messaging function

There are significant Parliamentary and civil society concerns about the categorisation threshold regulations, as laid by the Government and the advice from Ofcom on which they are based, as they do not use the full flexibility afforded by the OSA to address small but risky platforms. Please see our [background briefing](#) for more.

The OSA imposes additional duties on categorised services including:

- Fraudulent ads provisions (Category 1 and 2A)
- Transparency reporting (Cat 1, 2A and 2B)
- Publishing summaries of risk assessments (Cat 1 and 2A)
- Protection of democratic content, journalistic importance & news publisher content (Cat 1 only)
- User empowerment tools (Category 1 only)
- Enforcement of ToS requirements (Category 1 only)
- User ID verification (Category 1 only)
- Additional freedom of expression rules (Cat 1 only)

Please [see our table](#) for details on which duties fall on which category of service.

Implementation status: the regulations to set the categorisation thresholds have not (at the time of writing) been approved by Parliament. Once they have been approved, Ofcom will publish its register of categorised services and its final transparency reporting guidance, on whichb [it consulted](#) last year. Consultations on the additional duties for categorised services are likely to follow in 2026.

In force?: no, Parliamentary approval is still required for the regulations. The first duties (on transparency reporting) will not come into force until the register of categorised services is published along with the final guidance.

¹ The House of Lords will debate the regulations on 24 February; Lord Clement-Jones has tabled a regret motion. Please see [our briefing](#) on the background to this issue.

3. The Online Safety Act: enforcement and redress

a) [Enforcement](#)

The OSA gives Ofcom a [series of enforcement powers](#). These include:

- Issuing notice of contravention (provisional notices and confirmation decisions)
- Issuing penalty notices - these can be up to £18m or 10% of global revenue, whichever is higher
- Business disruption measures (service restriction orders and access restriction orders)

Ofcom also has significant information-gathering powers which it can use to inform its supervision and enforcement work (including the possibility of skilled persons' reports). Those who do not comply with an information notice commit an offence. Supervision of the largest companies has been undertaken by the regulator since it received its initial powers under Royal Assent; a taskforce to monitor small, but risky services has also been set up.

Please see [our explainer](#) for more details of the enforcement powers.

Implementation status: Ofcom has consulted on its information-gathering and enforcement approach. Its [final guidance](#) on information gathering was published in February 2025. Enforcement of the regime [has begun](#), focusing first on compliance with the illegal harms risk assessment duties.

b) [Redress](#)

There is no individual user right of action under the OSA and Ofcom is not set up to deal with individual complaints (contrast the broadcasting regime). The OSA establishes a supercomplaint regime, where designated bodies can raise concerns about systemic or market wide problems. DSIT is responsible for setting up the designation framework for bodies that are eligible to take supercomplaints. Although it [consulted on this framework](#) in early 2024, the regulations have not yet been laid; they are promised in "spring 2025".

The Act does not affect the possibility of a person directly suing another who has posted or shared content in contravention of the first person's rights. The Act does not change the principle of intermediary immunity from liability in such circumstances.

4. The Online Safety Act: implementation timetable

- By 16 March 2025: illegal harms duties in force. All providers must complete an illegal harms risk assessment and from this date (when codes are expected to be in force) must take measures to ensure compliance with them.
- By 16 April 2025: all providers must have completed a children's access assessment (under-18s).
- April 2025: Ofcom publishes guidance for children's risk assessments and its codes of practice.
- April 2025: Ofcom publishes a further consultation on the next iteration of the illegal harms codes of practice
- By July 2025: child safety duties in force. Services likely to be accessed by children must have completed a "suitable and sufficient" children's risk assessment and implement the safety measures required (including age assurance).
- Spring 2025: regulations for super-complaints designation regime (DSIT)
- Spring 2025: final Statement of Strategic Priorities for Online Safety (DSIT)
- Summer 2025: Ofcom publish register of categorised services (subject to regulations receiving Parliamentary approval)
- Summer 2025: Ofcom report on researcher access to data; DSIT consultation on regulations to bring researcher access regime provisions, included in Data (Use and Access) Bill, into force.
- Summer/winter 2025: draft and final transparency notices issued to categorised services
- Early 2026: consultation on additional duties on categorised services
- TBC 2026: revised illegal harms codes of practice

Further detail is provided in the latest version of [Ofcom's roadmap](#) and in our [detailed table](#).

5. The Online Safety Act: gaps and next steps

There are a number of civil society concerns about the approach that Ofcom has taken to implementation, which are detailed on our website. (See for example, [our joint response](#) to the illegal harms draft codes and [our concerns](#) about Ofcom's approach to engagement and consultation; our explainer on [safety by design](#), and [our joint response](#) on how the weaknesses in the approach to both sets of codes of practice proposals will mean that the guidance on protecting women and girls, on which Ofcom [is consulting](#), will not deliver what Parliament intended.)

Many of these issues stem from Ofcom's narrow interpretation of the Act so the Government has routes, if it wished, to clarify or tighten up the requirements to assist with Ofcom's approach. **We have provided a list of such targeted amendments in the annex.** The final version of the Government's [Statement of Strategic Priorities for Online Safety](#) is due to be published shortly.

Other gaps have also emerged which require more substantive policy development, consultation and legislative action - either through amendments to the OSA or separate legislation. These include:

- Concerns over the limits of the OSA to address the spread of online hate and related mis/disinformation that was evidenced during the Southport riots last summer (see our explainer [here](#)). The Government indicated that it might review the OSA but no further details have been provided on this.
- Addressing the UK's [epistemic security](#), particularly in the light of hostile activities from Elon Musk and the recent downgrading of Meta's user protections.
- Tightening up protections for children, particularly in light of the parent-led campaigns about smartphone use and screentime; Josh MacAlister's PMB, due for introduction in early March, addresses some of these issues. (See our [recent blog](#) on where the protections for children have got to since the work on the OSA began.)

March 2025

About the Online Safety Act Network

The Network was formed in November 2023 to coincide with the [Online Safety Act 2023](#) (OSA) receiving Royal Assent. It continues the work undertaken by [Carnegie UK](#) during the passage of the Online Safety Bill through Parliament, supporting civil society organisations, policymakers and Parliamentarians of all parties with expert analysis, briefing and legislative support to advance their various harm reduction interests and providing a regular space for collaboration.

We currently have over 70 civil society organisations represented in the network - from some of the largest children's NGOs and research institutions to small, grassroots charities - as well as individual academics and campaigners. We work constructively with representatives from the Government, Ofcom and other regulators, and members of both Houses of Parliament.

The Network provides:

- [Expert interpretation, analysis and thought leadership](#) on forthcoming consultations, secondary legislation and research to aid Network members' individual or joint activities and submissions.
- Regular opportunities for those with an ongoing interest in the successful implementation of the OSA to connect and collaborate.
- A curated [repository of relevant information](#) and reference material related to the OSA, its development and its current status.
- A [fortnightly newsletter](#) covering the latest online safety news in the UK and elsewhere.

ANNEX: Proposed technical amendments to the OSA 2023 to address emerging gaps in implementation

1. Introduce a general obligation on services to take reasonable steps to mitigate all the risks identified in their risk assessment.
2. Introduce minimum standards of terms of service for category 1 social media and search, e.g. for Equality Act protected characteristics, to provide a baseline of protections for users in the UK.
3. Insert a “no rolling back” clause to maintain ToS protections for users in the UK as they were at Royal Assent.
4. Introduce a requirement for Ofcom to produce a code of practice on safety by design, to deliver the objective set out in section 1 (3) and to focus more on harm caused by features and functionalities. This would underpin the existing, largely content-focused codes.
5. Remove the requirement in Schedule 4 for measures to be “clear and detailed”, which is contributing to Ofcom’s high evidential threshold and limiting the scope of the codes.
6. Remove the “safe harbour” provisions relating to the code of practice (section 49(1)) which means that companies can be in compliance with their safety duties without addressing all the risks they have identified on their service.
7. Upgrade the VAWG guidance to a code of practice to make it enforceable.
8. Amend the categorisation regulations to ensure the intent of the Act - that category 1 includes small, risky platforms - is delivered.
9. Review the list of priority offences to ensure that self-harm offences have parity with suicide.
10. Clarify the position on private messaging and end-to-end encryption, to ensure there are no safe havens for illegal content, such as child sexual abuse material
11. Set a minimum age limit for children’s access and tighten up the wording in the Act around the requirement for providers to deliver age-appropriate experiences for children to ensure Ofcom delivers measures to bring this into force.

Additional legislative amendments

12. Strengthen the researcher access to data provisions in the Data (Use and Access) Bill
13. Strengthen the NCII Offence introduced in the Data (Use and Access) Bill to ensure recourse and/or civil redress for individuals; and make NCII content illegal.