

THE ONLINE SAFETY BILL: TECHNICAL IMPLICATIONS

Internet Society England
10 June 2021



WHO WE ARE

- The UK's largest grassroots digital rights campaigning organisation
- Funded by grants, corporate support, and over 3,000 paying supporters and members
- 15 years of campaigning
- Areas of focus: online privacy, online freedom of expression, and online surveillance
- We collaborate with the Internet Society on legislative issues which impact the application/service layer



**“MAKING THE UK THE SAFEST PLACE
IN THE WORLD TO BE ONLINE”**

Draft Online
Safety Bill

1. Content moderation and service design duties

Duty	User to user	Search	All services	Likely Children*	Category 1	Enforceable by Ofcom	Citation**
Illegal content risk assessment	✓	✓	✓	✓	✓	✓	2 / 2 / 7 2 / 3 / 19
Children's risk assessment	✓	✓	✓	✓	✓	✓	2 / 2 / 7 (9) 2 / 3 / 19 (4)
Adults risk assessment	✓				✓	✓	2 / 2 / 11
Illegal content duties	✓	✓	✓	✓	✓	✓	2 / 2 / 9 2 / 3 / 21
Freedom of expression and privacy duties	✓	✓	✓		✓	✓	2 / 2 / 12
Democratic content duties	✓				✓	✓	2 / 2 / 13
Journalistic content duties	✓				✓	✓	2 / 2 / 14
Reporting and redress duties	✓	✓	✓	✓	✓	✓	2 / 2 / 15
Record keeping and review duties	✓	✓	✓	✓	✓	✓	2 / 2 / 16
Duties to carry out risk assessments	✓	✓	✓	✓	✓	✓	2 / 3 / 17
Safety duties for services likely to be accessed by children	✓	✓		✓		✓	2 / 2 / 10 2 / 3 / 22
Assessments about access by children	✓	✓	✓	✓	✓	✓	2 / 4
Transparency reports	✓	✓			✓	✓	3 / 1 / 49



2. Risk assessment requirements (2 / 2 / 7)

Requirement to identify, assess, and understand:	Illegal content risk assessment	Children's risk assessment	Adults risk assessment
The user base	✓	✓	✓
Risk to users of encountering illegal content (Terror / CSEA)	✓		
Level of harm to users of illegal content	✓		
The number of children accessing the service by age group		✓	
Level of risk to children of encountering each kind of priority primary content		✓	
Each kind of priority primary content that is harmful to children		✓	
Each kind of primary content that is harmful to children or adults, with each one separately assessed		✓	✓
Non-designated content that is harmful to children		✓	
Level of risk of harm presented by different descriptions of content that is harmful; for children by age group		✓	✓
Level of risk of functionalities allowing users to search for other members including children		✓	
Level of risk of functionalities allowing users to contact other users including children		✓	
Level of risk to adults of encountering other content that is harmful			✓
Level of risk of functionalities of the service facilitating the	✓	✓	✓

Screenshot



presence or dissemination of illegal content, identifying and assessing those functionalities that present higher levels of risk			
The different ways in which the service is used, and the impact that has on the level of risk of harm that might be suffered by individuals;	✓	✓	✓
Nature, and severity, of the harm that might be suffered by individuals by the above, including children by age group	✓	✓	✓
How the design and operation of the service (including the business model, governance and other systems and processes) may reduce or increase the risks identified.	✓	✓	✓



3. Administrative duties

Duty	User to user	Search	All services	Likely Children*	Category 1	Enforceable by Ofcom	Citation**
Notify Ofcom for fee payments	✓	✓	✓			✓	3 / 1 / 51
Information notices	✓	✓	✓			✓	4 / 5 / 70
Designation of a person to prepare a report	✓	✓	✓			✓	4 / 5 / 74 (4)
Assist the person preparing the report	✓	✓	✓			✓	4 / 5 / 74 (8)
Cooperate with an Ofcom investigation	✓	✓	✓			✓	4 / 5 / 75 (1)
Attend an interview with Ofcom	✓	✓	✓			✓	4 / 5 / 76 (2)
Duty to make a public statement	✓	✓	✓			✓	6 / 112 (3)
Information in connection with services presenting a threat	✓	✓	✓			✓	6 / 112 (5)



HOW WILL IT CHANGE INTERMEDIARY LIABILITY?

Draft Online
Safety Bill

THE TRADITIONAL INTERMEDIARY LIABILITY MODEL

Responsibilities:

- Due diligence
- Transparency of T&Cs
- Respecting user freedoms
- No general monitoring obligation

Scope:

- Objectively illegal content
- Moderation by moderators
- Good Samaritan provisions for moderation
- Content

Due process:

- Consistent treatment
- Conditional immunity for liability
- Notice and action (takedown)
- Regulation by courts



A “WORLD LEADING” INTERMEDIARY LIABILITY REGIME?

Responsibilities:

- Duty of care obligation
- Wrapping <18s in cotton wool
- Policing user behaviour
- A general monitoring obligation

Scope:

- Both illegal and legal but “subjectively harmful” content
- Moderation and modification by AI/ML/“safety tech”
- Content, conduct, and contact

Due process:

- Two-tier system of obligations
- Action then notice
- Regulation by regulators
- Business disruption measures



PRIVATE MESSAGING

Draft Online
Safety Bill



PRIVATE MESSAGING IN SCOPE

- The prohibition of end-to-end encryption, and the criminalisation of its use, is a singular goal of government
- Recent messaging strategies by government have openly sought to associate e2e with child abuse
- OSB will make private messages subject to the Duty of Care on illegal content (terrorism & CSAM)
- The Code of Practice on private messaging is likely to include the prohibition on end-to-end encryption to facilitate scanning of our messages for T & CSAM



ONCE THE DOOR IS OPEN TO OUR PRIVATE MESSAGES...

- It is not outside the scope of reason to expect the Duty of Care to demand the scanning of our private messages for **legal and subjective content** in addition to illegal content
- Companies will have no choice but to do this if they want to remain within the law
- Safety tech providers, supported by government, are seeking to erode privacy and computer misuse laws to facilitate proactive content scanning and modification
- One think tank has proposed that the Duty of Care should be used to impose criminal charges **retroactively** onto developers who are **currently** using e2e encryption – with no exceptions for e.g. financial transaction security



MANDATORY AGE AND IDENTITY VERIFICATION

Draft Online
Safety Bill



AGE GATING BY THE BACK DOOR

- The Bill requires companies in scope to assess whether it is possible for children to access the service
- Companies can only declare that children do not access their service if they can prove they identify and verify child users through technical systems or processes, such as age verification or assurance
- In other words, if you are not age gating your service, you are in violation of the Duty of Care and the Bill.



HOW IS THIS DIFFERENT?

- Past debates on age verification, such as the Digital Economy Act, concerned access to explicit adult content
- The OSB expands age verification from adult material to **all content and interactions, across all sites and services**, regardless of scope or proportionality, for anyone accessing any site or service based in the UK
- Aggressive corporate lobbying from age verification providers who want government to legislate in favour of their business model is now expanding the scope from age verification to identity and nationality verification



YOUR PAPERS, PLEASE

Type of user activity	Level of verification required	Rationale
View universally appropriate content.	No verification required.	Low risk of harm.
View age-sensitive content.	Verification of age attribute.	Prevent underage access of adult material.
Sharing universally accessible content	Verification of full name, nationality attributes.	Accountability for uploading content and for any content rating.
Sharing age-restricted content.	Verification of age, full name, nationality attributes.	Accountability for content/comments and content rating.
Universally accessible interaction with other users.	Verification of age, full name, nationality attributes.	Accountability for interaction. Prevent inappropriate adult/child contact.

BUSINESS DISRUPTION MEASURES

Draft Online
Safety Bill



BUSINESS DISRUPTION MEANS GOVERNMENT SHUTDOWNS

- Provisional notice of enforcement action for either illegal content (Terrorism and CSAM), or any one of the 25 subjective duties on both illegal and legal content
- Penalties enforceable to a group of entities
- Penalties for failure to comply with a technology notice
- Service restriction orders to providers of ancillary services (payment gateways, etc) to stop supporting a service
- Access restriction orders to block UK access to a site or service at ISP/network/app store level



HOW WILL THIS WORK IN PRACTICE?

- The Secretary of State for DCMS will need to go through both Ofcom and the courts to issue a business disruption measure...
- ...unless it is a matter of national security or interest, in which case no process or advance warning will be required
- The Bill therefore sets the stage for government to use any number of issues or public concerns as leverage to block platforms, services, and messaging apps from the UK



A BRITISH WEB FOR BRITISH PEOPLE

Draft Online
Safety Bill



WHAT DOES THAT ALL MEAN?

- Government has long said the goal of the OSB is to “make the UK the safest place in the world to be online”
- **It is now clear beyond any doubt that this means a British splinternet**, with the UK web – and access to it – subject to multiple speech and behavioural monitoring filters, verification gates, and disruption measures, at the technical levels



WHAT DOES THAT ALL MEAN?

- The responsibilities for this, and the compliance costs, will be outsourced to companies in scope of the regulation, who will also pay fees to the regulator
- Not only is government fine with this - they are also positioning these compliance obligations as a lucrative business opportunity



WHAT'S THE GOSSIP?

Well....



- Pre-legislative scrutiny committee begins work this month
- MPs view this Bill solely through the lens of
 - 1) social media
 - 2) abuse towards themselves and constituents
- MPs are being heavily lobbied by corporate interests and by celebrities
- There is a split across Whitehall
- There is a split across the Conservative Party
- This is going to run on and on and on and on and...



SETTLE IN, FOLKS.