

A student led computer science society dedicated to exploring computing, technology, and the ideas shaping our digital world.

Von Neumann Society - CCGS

Should social media companies be responsible for harmful content on their platforms?

Research Report by Ashwin Muralitharan

Who receives the blame for harmful content on social media?

A teenager in the UK can open TikTok and be shown videos picked for them by an algorithm, not by anyone they actually follow. Most of what turns up is harmless. Some of it isn't. For a long time, if something harmful did appear, the blame went to the person who posted it and the platform was mostly left alone. That has now changed, and the Online Safety Act 2023 makes platforms legally responsible for a lot more than before. In this report I will look at how UK law got to this point and whether it is the right approach.

How the law used to work:

Until recently, the main rule was a defence in the Electronic Commerce Regulations 2002. A company that hosted content posted by users was not liable for it as long as it did not know the content was illegal, or, once it found out, it removed it quickly (Infolaw). That made sense in 2002, when nobody expected a website to read every comment before it went up.

The problem was that the law never said how quick "quickly" had to be, and it was unclear whether actively checking content might count as knowing about it. In other words, a platform had very little reason to look closely at what was on its site.

Why did it change?

A big turning point was the death of Molly Russell, who was 14. In 2022 a coroner ruled that social media was a factor in her death, after she had viewed content about suicide, self-harm and depression (Mind). Her case made a lot of people ask whether platforms should really be treated as bystanders. Charities such as Mind then pushed the government to bring the Online Safety Bill back to Parliament, and it became law in October 2023.



A student led computer science society dedicated to exploring computing, technology, and the ideas shaping our digital world.

Von Neumann Society - CCGS

The Online Safety Act 2023:

The Act has been switched on in stages. Platforms had to start dealing with illegal content in March 2025, and the duty to protect children began in July 2025 (GOV.UK). The main thing to understand is that it is about systems and processes. A platform has to work out the risks on its service, put measures in place to reduce them, and take down illegal content when it appears. It is not automatically punished for every single bad post.

The regulator, Ofcom, can fine a company up to £18 million or 10% of its qualifying worldwide revenue, whichever is greater, and in the most serious cases it can ask a court to block a service (GOV.UK). There is also a personal side. If a company ignores an Ofcom decision about children's safety, and this happened with the consent, connivance or neglect of a senior manager, that manager can face up to two years in prison (RPC).

Is it working so far?

It is still early, but Ofcom has started handing out fines. In 2026 the website 4chan was fined £450,000 for not having proper age checks to stop children seeing adult content. It was also fined for not carrying out an illegal content risk assessment and for having unclear terms of service, which came to £520,000 in total (SCL).None of these fines were for one particular post, they were for not having the right safeguards in place.

Can a platform be held responsible without being blamed for every single post?

The case for responsibility really comes down to who is able to do something about it. A victim cannot track down every copy of a harmful video, and Ofcom cannot read billions of posts. The platform, on the other hand, decides what gets recommended and makes money from people's attention. If it costs them nothing when things go wrong, it is hard to see why they would spend much on safety. The 2002 rules are a good example of this, because they gave platforms a reason to stay out of the way.

However, the biggest case against responsibility is the worry regarding free speech. The Open Rights Group warned that if platforms face huge fines or personal liability, they may delete large amounts of legal content just to be safe, which it calls "collateral censorship" (Open Rights Group, 2021). It also pointed out that the rules apply to companies of any size with UK users, which is a heavy burden for a small website that does not have a team of lawyers.

There is also a bit of a mismatch so far. From what I could find, the big fines so far have landed on smaller websites like 4chan, not on the giant platforms that most teenagers actually use. It will be a bigger test when Ofcom takes on one of those.

The supporter's view is that platforms profit from and control what is shown, so they should be responsible for the harm their systems cause. In contrast, the critics' view is that heavy penalties encourage over-deleting and hurt small sites, so the cost to free speech may be bigger than the benefit.

Conclusion:

I think the UK has taken the right direction. The old hosting defence let platforms stay out of the way, and the Online Safety Act at least makes them do some work to reduce harm. It also avoids the worst version of the idea, where a company is blamed for every post, by focusing on systems instead. However, there are real risks to free speech and to small websites, and the Act is only a year or two old, so we do not yet know whether it actually makes people safer. I have also mainly relied on law firm summaries and news reports rather than long-term studies, so this is something to look at again in a few years.