

Written evidence submitted by Legal to Say, Legal to Type

DCMS sub-committee on Online Harms and Disinformation

About Legal to Say, Legal to Type

‘Legal to Say. Legal to Type.’ is a coalition of civil society organisations and industry groups formed in response to the publication of the Draft Online Safety Bill. The campaign is calling on the government to amend the Bill and ensure the internet is kept “free, open and secure”.

Summary of Key Points

- The bill’s aim of creating a safer online environment free from illegal content is welcomed, but this bill will actually make it harder for law enforcement to properly hold online abusers accountable and will be catastrophic for freedom of expression.
- The bill will create two tiers of free speech online: free speech for journalists and politicians, and censorship for ordinary citizens.
- The bill's introduction of the “Duty of Care model” is overly simplistic. The lack of an explicit definition of the term “harmful” will effectively outsource decision making on what is and isn’t permitted to say from the UK parliament and courts to Silicon Valley tech companies. The threat of large fines will create a commercial incentive to over-censor.
- The bill will make it harder for law enforcement to properly hold online abusers accountable, as it forces platforms to delete valuable evidence before the victims of targeted harassment or threats to kill can ensure it is reported to the police.
- Racism, homophobia, and threats of violence are already illegal. But data shows that when they happen online it is ignored by authorities. After the system for flagging online hate crime was underused by the police, the Home Office stopped including these figures in their annual report all together¹. The bill should focus on this illegal content rather than empowering the censorship of legal speech, including additional funding and training for police forces. Where there are gaps in the law, parliament should legislate to make specific areas of speech illegal.
- This bill sets a dangerous international precedent and could lead to a more authoritarian clampdown on use of legal speech online.

¹ [UK stopped analyzing online hate data after police ‘underused’ system.. Politico, 2021](#)

Q. How has the shifting focus between ‘online harms’ and ‘online safety’ influenced the development of the new regime and draft Bill?

The shift from harms to safety has led to an expansion of the remit of this bill in a way that is catastrophic for freedom of speech.

The House of Lords’ Communication & Digital Committee were correct in their view that the government’s stance on harms and safety “is not the right approach”². The lack of definitions has outsourced decision making to Silicon Valley, and the threat of large fines will lead to over-censorship.

The government must simply focus on illegal content. There is no doubt racism, homophobia, and transphobia are harmful and have no place in our society, but they are already illegal. We do not need to create any new legislation for these offences. Where there are gaps in the law, parliament should step in and address them

Leaving room for ambiguity and vagueness will create a multitude of problems and lead to a less safe online environment.

Q. Is it necessary to have an explicit definition and process for determining harm to children and adults in the Online Safety Bill, and what should it be?

Yes. The lack of an explicit definition of the term “harmful” will effectively outsource decision making on what is and isn’t permitted to say online from the UK parliament and courts to Silicon Valley tech companies. The threat of large fines will create a commercial incentive to over-censor. The government must make clear what is illegal online and address any gaps via proper parliamentary procedure

This is especially worrying for marginalised communities who are already over censored online.

In 2018, when Tumblr implemented a blanket ban of all adult content, its system repeatedly misclassified innocuous LGBTQ+ content as inappropriate and removed it. There have also been examples of platforms exhibiting racial biases: an image-detection algorithm was found to be cropping out black faces in favor of white faces³. A 2019 study showed that leading AI models for processing hate speech were one-and-a-half times more likely to flag tweets as offensive or hateful when they were written by African Americans, and 2.2 times more likely to flag tweets written in African American English.

There is also a risk that the current bill would make it harder to hold abusers to account. Social media content has become a vital tool in identifying and prosecuting criminals, and if platforms are incentivised to remove content as quickly as possible, crucial evidence will be removed before it can be reported.

² [Free for all? Freedom of expression in the digital age](#) (Lords, 2021)

³ [Twitter apologises for 'racist' image-cropping algorithm](#). The Guardian, 2020

This bill offers protections to trolls and abusers, making them feel safer to abuse people online; they should be prosecuted, not simply deleted.

The “duty of care” principle is not sufficient and should either be removed or reframed around illegal content - with protections for marginalised communities in accordance with UK Equality legislation. The government must make clear what is illegal online and address any gaps via proper parliamentary procedure. Simply put, if it is legal to say, it should be legal to type.

Q. Does the draft Bill focus enough on the ways tech companies could be encouraged to consider safety and/or the risk of harm in platform design and the systems and processes that they put in place?

No. Relying on systems and processes without properly factoring in specific protections for ordinary people - and especially marginalised groups like the LGBTQ community - is the wrong approach. These systems must be designed with equalities legislation in mind. Failure to do so will lead to a divisive online environment with two tiers of free speech, where some are censored and others are free to speak their mind.

Tech companies will be in charge of what we say online, and excessive fines will bring a financial incentive to over-censor our online lives. In addition, hate groups will use the threat of these fines to pressure companies to remove content they do not agree with, as has already happened in the case of a LinkedIn removing a user’s coming-out post following complaints⁴.

Algorithms can create a more hostile and discriminatory online environment: for example, the way in which content is moderated means that it is twice as likely to be deleted if it is in Arabic or Urdu than if it is in English. Sites like Tumblr have opted for a full-scale ban on all “not suitable” content; the removal of which is powered by AI-driven content recognition. The effect of this decision has been to eradicate a platform that has been central to the LGBTQ+ community for self-expression and development⁵.

Q. What are the key omissions to the draft Bill, such as a general safety duty or powers to deal with urgent security threats, and (how) could they be practically included without compromising rights such as freedom of expression?

Urgent safety threats are always best handled by the police, not Silicon Valley. Funding and resources must be provided to law enforcement to help tackle egregious illegal content online.

Given this, it is obvious that there should be an expansion of legislation governing egregious crimes to include the developing online space - which should be policed by traditional law enforcement. To

⁴ [Brave teen came out to classmates by coming out in a dress for his prom](#). Metro, 2021

⁵ [Why Tumblr’s ban on adult content is bad for LGBTQ youth](#). The Conversation, 2018

facilitate this, resources must be provided to law enforcement so that they can effectively police the online space.

In 2016, research conducted by Middlesex University found that only 1 in 6 police officers who had investigated online grooming or indecent images of children online had been given any “specific” training⁶, and after the system for flagging online hate crime was underused by the police, the Home Office stopped including these figures in their annual report all together⁷. Rather than applying a new regulatory or bureaucratic approach to these issues - via targeting social media and big tech - the government should instead provide law enforcement with the necessary tools and resources to tackle this behaviour.

Q. Are there any contested inclusions, tensions or contradictions in the draft Bill that need to be more carefully considered before the final Bill is put to Parliament?

The bill's introduction of the “Duty of Care model” is overly simplistic. The lack of definition for the term “harmful” will effectively outsource decision making on what is and isn’t permitted to say online from the UK parliament to Silicon Valley tech companies. The threat of large fines will create a commercial incentive to over-censor. By removing evidence of harmful or illegal behaviour, tech companies will actually do what they are trying to prevent: they will make the online space less safe by protecting abusers from punishment.

To make these platforms truly safer, resources and funding must be directed towards law enforcement and education - to tackle the root causes of this behaviour - not regulatory bodies or bureaucrats. By failing to do so, this bill will actually make it harder for law enforcement to properly hold online abusers to account.

In addition, the government’s intention to hand regulatory oversight to Ofcom - already the adjudicator for other forms of communication - would be an unprecedented consolidation of power. Given that the Secretary of State has the authority to identify any additional content as “harmful” under the “priority content” classification⁸, this will enable the government to politicise content moderation and censor the opposition.

Under the “journalistic content” and content of “democratic importance” exemptions, this bill will censor ordinary citizens and allow journalists and politicians to speak freely online. This is a further outsourcing of our rights to Silicon Valley, who - given the lack of clear definition for “journalism” - will decide what is journalism and what is not. This sets a dangerous precedent and will censor countless citizen journalists and online news sources. The government must address this gap and make clear how non-traditional news sources will be handled under this bill.

⁶ [Enhancing Police and Industry Practice](#) (Middlesex University, 2016).

⁷ [UK stopped analyzing online hate data after police ‘underused’ system](#). Politico, 2021.

⁸ [Draft Online Safety Bill](#)

This bill has further exemptions for “democratically important content”: defined as “content promoting or opposing a political party”, yet it does not state to what degree this exemption will apply. Far-right or extremist public figures like Tommy Robinson have been banned from social media in the past for breaching hate-speech regulations, yet given he has previously ran in elections, it is unclear whether this bill would exempt him on political grounds. Failure to properly define who is covered by this exemption will mean people like Robinson, or former politicians and activists could be given special privileges online. This further divides society by creating one rule for politicians, public figures, think-tanks, and public bodies, and one rule for everyone else.

Q. What are the lessons that the Government should learn when directly comparing the draft Bill to existing and proposed legislation around the world?

The bill is a threat to our fundamental right to free speech. In June this year the French Constitutional Council found large parts of France’s online hate speech legislation unconstitutional⁹. It cited a breach of the legality test for “impermissible vagueness”: users must know with reasonable certainty whether they will be legally obliged to delete content they are about to post. This proposed bill - through its poor handling of what is legal or illegal online - threatens to do the same.

The penalties for failing to remove harmful content are too high and will result in a system of over-censorship and over-regulation. For comparison, Germany’s [Network Enforcement Act](#), has a maximum potential fine of £20 million. Compare this to the proposed fines in this bill of 10% annual turnover; for Facebook this would mean an £8.8 billion fine. The incentive then will be to over-regulate and over-censor to avoid these excessive fines.

This bill sets a dangerous precedent for further regulation and curtailing of our freedom of expression. In Pakistan, internet regulation began with the removal of “offensive” and “blasphemous” content in 2006-2008¹⁰. What followed was a government-led content filtering and website-blocking system. This included network monitoring and a country-wide YouTube ban in 2012 which took 3 years to lift¹¹. Without amending the problem areas of this bill, this could happen in the UK.

The UK government must recognise its power in the global community. The Pakistani government have published their Citizens Protection (against Online Harms) Rules¹², which has lifted the UK’s definition of “extremism” from its “Prevent” strategy and been used to clamp down on social media. This definition has been criticised by the both the UN Special Rapporteurs on the rights to freedom of peaceful assembly and association and on counter-terrorism¹³ for being “inherently flawed”. The UK must realise its power in the global community and be cognizant of other countries looking to it for guidance on these matters.

⁹ [French constitutional court blocks large portion of online hate speech law](#) (Euronews, 2020)

¹⁰ [The Anatomy of Web Censorship in Pakistan](#), (Nabi, 2013)

¹¹ [Pakistan lifts YouTube ban after three years](#) CNBC, 2016

¹² [Government of Pakistan \(Information Technology and Telecommunications Division\) Ministry of Information Technology and Telecommunication\): Citizens Protection Rules](#)

¹³ [Pakistan: Online Harms Rules violate freedom of expression](#) (Article 16, 2020)