

## **Online Safety Bill – Parliamentary Briefing from Microsoft**

December 2022

Microsoft has long supported the aims of the Online Safety Bill. As a company, we have a history of working with governments, civil society, academia, technologists, and others in industry to promote online safety. When the draft Bill was published in May 2021, we raised several critical issues with Government. While much progress has been made following the legislative scrutiny process carried out by the Joint Committee on the Draft Online Safety Bill, as well as the recently announced decision to remove clauses focused on legal but harmful content for adults, we believe critical issues with the legislation still remain.

### **Issue 1 – The Bill would impact UK internet users’ access to information via search**

The Bill continues to include a range of obligations for search services, including a duty to use systems and processes designed to minimise the risk of users encountering priority illegal content via search, and child safety duties. From Microsoft’s perspective, there are several issues of concern here:

- It is not clear how search engines could ever operationalise the requirement to minimize the risk of users encountering priority illegal content and/or content that is harmful to children. The requirements fail to recognize the unique nature of search engines, which do not host content, as well as the important role search engines play in facilitating access to information.
- Without reviewing each of the 100 trillion+ indexable pages of the internet, it is impossible for a search service to know whether a given search result may link to illegal or harmful content.
- To fulfil these duties, search engines would be required to monitor all content across the indexable web using machine learning technology. There is no failsafe detection technology for most content types – whether illegal or harmful, making deployment of current technologies a blunt instrument. Substantial numbers of URLs linking to lawful and appropriate content would inevitably be included in these removals. Entire legal websites or services may no longer be identified through search.
- This is exacerbated by the broad range of content that may fall within the scope of the priority illegal offences identified in the Bill. Given the difficulty in making these judgments, including the need for context and even external evidence to determine intent, it is unclear how a search service could take meaningful measures to prevent access to content related to, for example, harassment, fraud, or illegal immigration without cutting off UK users access to lawful content, such as resources to support victims of stalking, or the provision of financial or immigration advice.
- Moreover, to meet the enhanced child safety age assurance requirements in the Bill, search engines would either have to assume all users are children and limit their access to search results that might be harmful to children, or, implement blanket age verification for all users to identify which are children and provide them a different search experience.
- Maintaining “safe search” for all users, including adults will negatively impact UK users’ access to information online, particularly related to some of the most important and sensitive topics facing Western civilization. The other alternative, universal age verification in order to age-gate users of search, would require all users to register and

be subject to data collection that is significantly intrusive to their personal privacy by requiring identity documentation. Age gating search will create a loss of anonymity for all UK users – many of whom use search to find information on deeply sensitive and personal issues, such as researching medical conditions, exploring their sexuality, or seeking advice on domestic violence situations. This would also disadvantage adult users who may not have the means to verify their identity.

- Regardless, these duties on search engines will result in the censorship of lawful opinions and make it more challenging for children of differing ages to access diverse information on a range of topics. Search services play a crucial, enabling role for the work of educators, researchers and journalists in this regard. To name an example, content related to the war in Ukraine or humanitarian conflicts could in some contexts be considered harmful. As such, video coverage of journalists or human rights organisations – which by some may be considered harmful but is essential from a freedom of information and press point of view - could end up being viewed as harmful content for children. The Bill, as it stands, would therefore inevitably chill freedom of access to information and other human rights.
- .
- Microsoft already has leading technology in place to minimise the risk of children’s exposure to harmful content, such as our SafeSearch tool. However, the Bill remains unclear in what is expected and we would encourage clarity in order to avoid the above issues, recognising the positive impact of search in the dissemination of knowledge and information. We urge Parliament to consider whether the list of priority illegal offences should be narrowed for search engines to avoid unintended consequences and to appropriately recognize the unique role they play in the internet ecosystem.

## **Issue 2 – Tech companies should not be expected to both determine and remove what is illegal beyond child sexual exploitation and abuse (CSEA) and terrorist imagery offences**

- Microsoft is fully committed to stamping out CSEA and terrorist content across our platforms. Microsoft already takes a range of measures to address such content, including the use of technical tools.
- However, we continue to be concerned that the Bill requires online service providers to make judgements on the legality or otherwise of a broad range of content. It will often not be clear on the face of a specific item of content (most particularly textual content), or from the limited information available, whether a UK court would decide that content was illegal. Such decisions can be difficult and involve a context-specific balancing of rights, which sits most appropriately with elected officials or the judiciary. These proposals risk inappropriately placing online service providers in the position of prosecution, defence, judge, and jury.
- Specifically, the Bill requires providers to determine the legality of content related to a wide range of priority offences, ranging from assisting illegal immigration, to harassment, to offenses involving drugs, firearms, and fraud, among other things.

Clause 166 says providers should consider all reasonably available contextual information and put in place systems that enable them to infer whether all the elements (including mental elements) of a UK legal offence are met. *Practical challenges*

- This would in theory require providers to train all moderators in highly complex areas of law and make “correct” judgments within minutes that in a court might be subject to months of debate.
- In almost all cases outside of the most obvious CSEA, it will be impossible to determine illegality on the face of a piece of content. Contextual information, background information about the provider of the content, as well as knowledge of the elements of each potential offence would be required. Deep contextual understanding of colloquialisms of individual subcultures may also be required to identify coded cues.
- In practice, such context and background information are unlikely to be available without some further intrusion into an individual’s privacy (to conduct what essentially becomes surveillance).
- Coupled with the wide variety of priority illegal content and the requirement to prevent users from encountering illegal content (see below), this is highly likely to result in an over-removal of content. Given the very wide array of content types in scope, this could have major impacts for UK users.
  - For instance, a post about illegal immigration is likely to contain very similar elements to a post about lawful processes to immigrate – a topic about which many people seek information and assistance in public forums. With such similarities in language, proactive technologies will not be able to reliably detect (and thus prevent) such content.
  - In other instances, language such as “I’ll kill you” may be flagged by automated content moderation as an illegal threat. But the statement might make complete sense if the user is engaged in a player vs player video game.
- The kind of decision-making required by the Bill would be challenging for a human moderator, let alone to appropriate training technical tools. These challenges are further exacerbated in relation to the Bill’s duties with regard to the facilitation of offences, which is a step further removed from potentially illegal conduct.
- UK users may also have legitimate concerns about the outsourcing of complex decisions on legality to private companies, many of which are not UK based. Such adjudication is a core function of different arms of government. Parliament should also consider the broader societal implications of a Secretary of State having the power to amend Schedule 7 to add further priority offences a. In this respect, Microsoft agrees with the suggestion of the Joint Committee on the Online Safety Bill who recommended that Government, in bringing forward future criminal offences, consults with Ofcom and the Joint Committee as to whether they should be designated as priority illegal offences in the legislation that creates them.

**Issue 3: It is unclear how the Bill will balance duties both to protect users and to uphold principles of freedom of expression**

The Government has recently reiterated its commitment to upholding the principle of freedom of speech, and made efforts in this regard via the removal of clauses focused on content that is legal but harmful for adults. However, there continues to be an inherent tension between the Bill’s obligations for providers to protect users from illegal content and content that is harmful to children, and the same providers’ duties to protect the rights of freedom of expression and privacy. As a result of this tension, the Bill does not provide sufficient clarity to support Ofcom in developing tailored, proportionate guidance that enables providers to

effectively meet their obligations across a diverse range of products and services. While the Bill focuses on risk-based approaches, what constitutes a proportionate approach to addressing those risks will differ, depending on the unique service and content type in question.

For the Bill to achieve its stated objective of safeguarding freedom of expression, this tension must be resolved through a recognition that tech companies should not be tasked with deciding what is illegal content. As per current process, the removal of illegal content should require a court order, building in an independent tribunal process that ensures that content is in fact illegal before requiring providers to remove it.

#### **Issue 4: More clarity is required on what is considered content that is harmful to children**

As it exists currently, the Bill imposes additional duties of care on certain user to user services and search providers with respect to “content that is harmful to children” Content is harmful to children if it falls within the categories that the Secretary of State will identify in secondary regulations, or if the content is “of a kind which presents a material risk of significant harm to an appreciable number of children in the United Kingdom” (Clause 54).

To meet the principle of legality, the Bill should ensure that any content deemed to be “primary priority content that is harmful to children” or “priority content that is harmful to children” is very clearly defined. To oblige providers to remove content based on non-specific definitions could unintentionally lead to the over-removal of content and, by extension, censorship and overreach – in clear opposition to one of the stated aims of the Bill: to safeguard freedom of expression online. To help build collective understanding, we also recommend greater specificity in the criteria that the Secretary of State must consider when developing regulations. For instance, no guidance is provided on what is meant by an “appreciable” number of children, nor on whether it is “appropriate” to apply the Bill’s duties to such content. To ensure any such regulations are developed based on solid evidentiary basis, it is important to have clarity around the procedures that OFCOM will need to apply, including consultation periods, independent review, etc.

Further clarity is also warranted with respect to “non-designated content that is harmful to children”. This is currently defined at clause 54(4)(c), and as with the other categories of content harmful to children, no guidance is provided on how to understand key elements, such as “material risk”, “significant harm” and “appreciable number”. Without such clarity, providers may be unable to meet the risk assessment and child safety duties related to such content.

#### **Issue 5: Conflict-of-laws issues must be resolved now, before it is too late**

The Bill is being developed in parallel to related legislation in other markets, including the EU, the USA and Australia. The European Union is in the process of updating the 2001 Electronic Commerce Directive in the form of the proposed [Digital Services Act](#), while policymakers in Australia and the United States, have offered their own laws or regulations governing content moderation. In many cases, these proposals take quite different approaches for promoting online safety to those set out in the Bill. In some cases, the Bill would impose diverging or even conflicting obligations on service providers to those set out in these other proposals.

Without aligning the Bill as closely as possible to these examples, unavoidable conflicts of laws may become apparent. Practical consequences for UK internet users may include reduced access to information, given the UK Bill currently includes incentives to over-moderate content. Equally, UK users may have less of an ability to express their views online than friends or family in other countries. UK citizens may also find that they have access to fewer online services, if some services opt to leave or reduce their presence in the UK market. Smaller UK companies or start-ups may also find themselves at a competitive disadvantage compared to companies based in other markets. With respect to Bing, users would likely have less access to information due to potential for over moderation of content.

The Joint Committee has recommended that Ofcom should have the power to share information and to co-operate with international regulators at its discretion. Microsoft urges Parliament looks to align the Bill as closely as possible to legislation across the UK's main economic partners. We also propose a conflict-of-laws provision which would excuse tech companies for not complying with the Bill in circumstances where doing so would force them to violate another law to which they are bound. We also urge Parliament to consider whether other specific protections may be appropriate, such as including a prohibition on general monitoring.