The Online Safety Bill and the role of technology in child protection

Summary

This paper argues that the Online Safety Bill (OSB) should not put its trust in emerging technology solutions to deliver child protection without rigorous analysis of their flaws, evaluation of the privacy trade-off, and a balancing emphasis on education and awareness.

The paper was prepared by BCS, The Chartered Institute for IT's Fellows Technical Advisory Group (F-TAG). It builds on existing positions around end-to-end-encryption (E2EE) and client-side scanning (CSS).

Further recommendations

Policy makers need improved understanding of the technology issues posed by the OSB, to avoid unintended consequences as the bill moves into law.

Legislation should address not just technical intervention, but also education, training of professionals, and public awareness programmes.

Online safety policies should aim to provide young people, their parents and advisers with an understanding of risks and how to mitigate them, rather than assuming technology will prevent those risks from arising.

Compromising end-to-end encryption (used by popular messaging apps) is not possible without introducing systemic risks and 'bugging' millions of users phones.

Age-verification proposals are not yet proven to adequately prevent illegal access to material. Similarly, current detection of child sexual abuse material (CSAM) is highly unreliable.

The UK's international reputation on data security is likely to decline with legislation that undermines encryption – as illustrated by public statements from WhatsApp and Signal.

Technology solutions will be imperfect and have unintended consequences

The recently published Women and Equalities Committee report "Attitudes towards women and girls in educational settings", highlighted the extent of sexual harassment in schools, with online pornography identified as a factor influencing these behaviours. The report called for the implementation of measures to prevent young people from accessing pornography, and BBC reporting on the findings, quotes the Department for Education as stating:*"Through the Online Safety Bill, technology firms will be required to enforce their age limits and protect children from being exposed to harmful material online".*

There is a history of government making technical proposals - for example, filtering by internet service providers (ISPs) and failed attempts for age verification - in past legislation to prevent young people from accessing pornography.

While technologists, and indeed young people themselves, have pointed out that bypassing UK specific age verification technologies is currently simple, there is still a view that technical intervention with age assurance on pornographic websites will prevent access.

Labour MP Sarah Champion has called for a review of virtual private networks (VPNs) by OFCOM and the need for government to "find solutions" to the use of VPNs to bypass geographically bound age-assurance technologies.

Effective prevention of children accessing pornography must be a priority for government and society. Yet technology cannot, of itself, prevent access to illegal material.

Yoti is a UK-based company providing token-based authentication services – which are undoubtedly effective, but which present their own challenges for many in the population who might not own a passport or driving license. Yoti is now making inroads into using machine learning to estimate an end user's age based in facial recognition. The company claimed in a white paper last year:

Our True Positive Rate (TPR) for 13-17 year olds being correctly estimated as under 23 is 99.65%. This gives regulators a very high level of confidence that nobody under age will be able to access adult content. Our TPR for 6-11 year olds being correctly estimated as under 13 is 98.91%. Our solution is configurable to meet any regulations that require prior consent before age estimation is used.

While those figures are impressive and undoubtedly show increases in the efficacy of machine learning-based estimation systems, they are not claiming they can correctly estimate ages under 18 but rather ages under 23. The other priority areas identified in the bill are terrorism, child sexual abuse material (CSAM) and grooming. It is obvious that terrorists and indeed organised criminals will have the attention and resources to simply not use commercial platforms for criminal messages, or that they will use VPNs to access alternative communications networks.

The technical countermeasures for different types of activity vary greatly, especially in their degrees of viability and intrusiveness. For example, the detection of grooming activities involves a high degree of semantic and contextual inference, making it unlikely that technical solutions will deliver workable or societally acceptable results, particularly at scale. For instance, if I say "deliver 20 litres of hydrogen peroxide", am I building a bomb or re-stocking my hairdressing salon? If I say "deliver 20kg of apples", am I re-stocking my fruit stall, or ordering plastic explosives?

Compromising E2EE is not possible without introducing systemic vulnerabilities and risks

BCS is committed to ensuring technology is beneficial for the public and has previously said that it is vital that end-to-end encryption, which allows secure private messaging, is not undermined by the scanning of personal communication. However, *"End-to-end encryption should not be rolled out without appropriate safety mitigations, for example, the ability to continue to detect known [CSAM] imagery", the government says.*

The Online Safety Bill and politicians still talk about "accredited technologies" being used to intercept encrypted messages without privacy implications.

A recent House of Lords amendment calls for a "skilled person" to "write a report" that assesses the privacy impact before such technologies are used. However, 70% of BCS members are not confident it was possible to have both truly secure encryption and the ability to check encrypted messages for criminal material. It is not yet clear what the qualifications and assessment criteria of that skilled person would be.

Analysis of the potential to develop such technologies using CSS or homomorphic encryption (the ability to perform computation on encrypted data that you can't decrypt) shows that at a technical level, any foreseeable accredited technology would compromise privacy.

It would have to be implemented as a "bug" on each smartphone, that would be connected to the government or a technology provider. The risks of introducing such a systemic vulnerability will likely outweigh the speculative benefit to law enforcement.

The privacy trade-off should be evidence-based

While the Internet Watch Foundation is reporting increases in CSAM online, it is also important to be aware that E2EE has grown significantly since 2015 and has not led to a decrease in UK prosecutions around images of abuse. This is supported by a study by the Max Planck Institute showing that increased digital surveillance in Germany did not increase convictions.

The trade-off for the privacy of all citizens (including children) must be evidence-based and proportionate.

The trade-off must also make clear that a technical approach effective against already known CSAM images would not be effective against the other types of criminality. The technology capable of detecting the redistribution of already known CSAM images cannot simply be repurposed to detect new CSAM images, grooming and terrorism, and it is unsafe to draft laws based on the assumption that it can.

The trade-off also needs to consider that successful disruption of networked criminal activity is possible even in the presence of fully encrypted communication systems. It has been shown that the police can conduct cross-border major operations and penetrate fully encrypted communications where they need to. For example Silk Road, Alpha Bay, Encrochat and Sky ECC.

On the other hand, researchers from Imperial College London say that we don't understand the risks yet well enough to ask for the deployment of CSS technology on hundreds of millions of devices. The UK's National Research Centre on Privacy, Harm Reduction and Adversarial Influence Online (REPHRAIN) agrees and called on politicians in Britain "to consider independent scientific evaluation before voting through the bill" in the *Open Letter from Security and Privacy Researchers in relation to the Online Safety Bill.*

Further attempts to compromise internet security will damage the UK's reputation for better regulatory approaches

WhatsApp, Signal and other encrypted communications platforms have seriously criticised the bill and even threatened to stop providing encryption for UK citizens rather than comply. Some platforms will be forced to either leave the UK market or undermine the security and privacy of all their users, including the most vulnerable.

It is also likely that the UK international reputation on data security and as an effective regulator of technology will suffer. As well as undermining the market for products developed in the UK (if they are known to be insecure), the OSB would make the UK the weak link in cross-border communication.

There is evidence of this from Australia's bill on Telecommunications "Assistance and Access" Act 2018 which led to a reduction in business investment. The researchers found that even the slightest distrust in data security can have long-lasting adverse effects on a country's economy.

Conclusion

In conclusion, BCS advocates for a comprehensive approach to online safety, highlighting the need to consider not only technical solutions but also education, awareness, and privacy concerns. BCS warns against compromising encryption and emphasises the importance of understanding the risks and implications of technology before enacting legislation.