Before It Begins

A Legislative Path to Prevent Livestreamed Child Sexual Abuse Through On-Device AI Safeguards

July 2025

Before It Begins

A Legislative Path to Prevent Livestreamed Child Sexual Abuse Through On-Device AI Safeguards

July 2025

1. Executive Summary

The United Kingdom is now among the top three global consumers of livestreamed child sexual abuse. This abuse occurs in real time, over encrypted video calls, often facilitated by impoverished caregivers in low-income countries and orchestrated by paying offenders in the UK and other high-income nations. These events are not hidden in the dark web—they are occurring on mainstream apps and platforms, through everyday devices.

Unlike existing child sexual abuse material (CSAM) in image or video form, livestreamed abuse is exceptionally difficult to detect or disrupt. There is no stored media to scan or remove. The violation occurs in the moment—and is gone. The only viable path to intervention is prevention at the point of access.

A new class of technology now offers this opportunity. On-device AI classifiers, embedded within operating systems of internet-connected, camera-enabled devices, can detect, block, and disrupt attempts to capture or view CSAM—including livestreamed abuse—before it happens. These systems operate entirely on-device, meaning they are compatible with end-to-end encryption and preserve user privacy.

This paper proposes that the UK Government act decisively to protect children by mandating the inclusion of such safeguards on all devices sold within the UK market. This would extend the intentions of the UK Online Safety Act beyond platforms and into the devices themselves—addressing abuse at its root, not merely its digital distribution.

Key benefits of this proposal:

- Prevention before harm: Blocks the creation and consumption of livestreamed abuse in real time. - Privacy-preserving: Detection occurs on-device; no content is uploaded or shared externally. - Encryption-compatible: Fully aligned with encrypted communications; no backdoors required. - Scalable and sustainable: Applies to all devices entering the UK market; builds long-term resilience.

The technology exists. The ethical case is clear. The legislative gap is identifiable and addressable. The UK now has the opportunity—and the responsibility—to act.

2. The Scale of Harm

Livestreamed child sexual abuse is not a distant or rare crime—it is a growing, technologically enabled market of real-time violence. It differs fundamentally from historic CSAM offences: there are no files to scan, no caches to flag, and often no digital trail. The abuse takes place live, over encrypted video calls, directed by the viewer, and disappears as quickly as it began. There is no afterimage. Only trauma.

According to the UK's National Crime Agency (NCA), the UK is the third largest global consumer of livestreamed child sexual abuse. Europol has declared it "the main form of commercial sexual exploitation of children" worldwide. In 2022, a joint study by International Justice Mission and the University of Nottingham Rights Lab estimated that nearly half a million Filipino children—approximately 1 in every 100—were trafficked to produce livestreamed or other CSAM content.

The cost to access these livestreamed abuses is often trivial—sometimes as little as £15 per session. The viewing offenders (known as the demand side) pay to direct the abuse from remote locations—often the UK, EU, Australia, or the US—while the acts are carried out by facilitators in low-income countries. These facilitators may be relatives, neighbours, or coerced intermediaries.

Research consistently shows that men who commit online child sexual offences are 2.5 times more likely to seek in-person contact with children. For livestream offenders, this risk may be even greater.

Nearly 1 in 10 men in the UK, US, and Australia report having engaged in some form of online sexual offending behaviour against children at some point in their lives. This is not a fringe issue. It is a systemic one. And it is growing.

At the 2024 Global Ministerial Conference on Ending Violence Against Children in Bogotá, the UK pledged to lead global action to prevent online child sexual exploitation. This paper now offers a concrete, implementable next step—aligned with that pledge, and urgently needed.

3. The Technological Opportunity

The challenge of detecting livestreamed child sexual abuse lies in its nature: it is ephemeral, encrypted, and often indistinguishable from legitimate use—until it is too late. Conventional content moderation systems, server-side scanning, and reporting mechanisms cannot intervene in time, especially on encrypted platforms.

But a breakthrough now exists.

On-device AI classifiers—machine learning models embedded within the operating systems of smartphones, tablets, and laptops—can detect high-risk behaviour patterns and illegal content before it is transmitted or viewed. This technology represents a new frontier: real-time prevention without surveillance.

Unlike server-based solutions, on-device classifiers operate locally: - They analyse camera input or outgoing/incoming media in real time. - Detection is handled entirely within
the device; no data is uploaded or shared externally. - When illegal content is detected, the system can block recording or transmission, trigger local warnings, or generate anonymous alerts.

This architecture protects both user privacy and child safety. Because the classifier operates entirely on the device, end-to-end encryption remains intact. There are no backdoors, no scanning of private conversations, and no compromise of civil liberties.

Some forms of this technology are already in use—major tech companies apply nudity-blocking AI models to children's accounts. However, livestreamed abuse most often originates from adult devices. What's needed is the extension of these protections to all internet-connected, camera-enabled devices, and the mandating of their inclusion at the OS level.

Modern classifiers, particularly those trained to detect prepubescent nudity, now operate with very high levels of accuracy and low false positives. When deployed with safeguards—such as local-only processing and narrow classifier scope—they offer an ethically sound, legally compatible solution to a global crisis.

4. Legal and Ethical Considerations

In the pursuit of child protection, particularly against such devastating and systemic abuse, governments must act with urgency—but also with care. Any intervention must be legally sound, ethically coherent, and publicly defensible. The proposed solution—on-device AI safeguards—is one of the few available tools that meets all three criteria.

Because detection happens entirely on-device, this approach is fully compatible with both: - UK privacy law, including the UK GDPR and the Human Rights Act (1998) - End-to-end encryption protocols used by platforms like WhatsApp and Signal

There is no central scanning, no government access to user data, and no interception of communications. Rather, detection remains within the device ecosystem—akin to parental control settings or app filters.

Ethical use of this technology hinges on: - Clear legislative boundaries specifying what content triggers intervention - Narrow classifier scope tuned to illegal content only - Independent oversight and third-party evaluation - Opt-in features for alerts or guardian notifications where appropriate

By targeting the point of production or viewing, this approach avoids blanket surveillance. It becomes a tool of preventative care—effective unless wrongdoing is attempted.

The UK's Online Safety Act (2023) is a powerful precedent. However, it applies only to platforms and not to devices or operating systems. Extending this scope would close a critical regulatory loophole and fulfil the UK's international pledge to lead in child protection.
Globally, jurisdictions like the EU and Australia are moving in this direction. The UK has the infrastructure and moral standing to lead.

5. Policy Recommendation

To prevent livestreamed child sexual abuse at scale, the UK Government must extend its regulatory scope beyond platforms and into the devices that enable abuse. This means legislating at the operating system and hardware level.

We propose the following:

1. Amend the Online Safety Act to: - Mandate on-device AI classifiers for internet-connected, camera-enabled devices sold in the UK. - Require classifiers to detect illegal CSAM, especially livestreamed content. - Ensure they operate as part of the OS, preserve privacy, and avoid weakening encryption.

2. Require manufacturer compliance: - All major device manufacturers must ensure compliant detection systems are embedded for UK sales.

3. Establish an independent certification body to: - Review classifier models - Certify compliance - Conduct audits and respond to user concerns

4. Provide research and development support: - Public grants and technical partnerships to accelerate model refinement - Implementation support for smaller tech firms

5. Lead international coordination: - Propose model legislation through Commonwealth, UN, G7, and digital safety coalitions - Share ethical standards and research tools

This is a chance to set a global standard in protecting children—through technology that is proactive, private, and preventative.

6. Implementation Pathways

The legislative framework must be paired with a clear and collaborative implementation strategy.

Phase 1: Regulatory Preparation (0–6 months) - Establish a government-led Working Group on Device-Level Child Safeguards - Begin drafting legal amendments - Launch public consultation

Phase 2: Standards and Certification (6–12 months) - Develop detection and privacy standards - Create a certification body to audit classifiers - Initiate pilot programmes with select manufacturers
Phase 3: Market Transition (12–24 months) - Legislate and enforce mandatory inclusion - Launch a public awareness campaign - Support small businesses in compliance

Phase 4: International Expansion (24–36 months) - Align efforts through global partnerships - Share findings and open technical frameworks - Publish public impact reports

Throughout all phases, transparency, accountability, and proportionality must guide implementation—ensuring protection does not come at the cost of dignity or rights.

7. Conclusion: Leadership Through Light

There are moments in policy where the choices before us are not abstract, but piercingly real. This is one such moment.

Livestreamed child sexual abuse is not theoretical. It is happening now—through UK devices, on UK networks, directed by UK offenders. It is facilitated by the infrastructure of modern life and suffered by children in silence.

We now possess the technology to stop abuse in real time—without violating privacy or breaking encryption. This paper outlines a lawful, ethical path to embed protection at the point of harm.

The UK has the infrastructure, international influence, and moral opportunity to lead. Let this be the moment we act—not reactively, but decisively. Not after the harm, but before it begins.

Not everything in this world can be prevented. But this can. Let it begin—before it begins.

8. References

1. UK National Crime Agency (NCA). (2024). The Rise of Livestreamed Child Abuse and Britain’s Role In It. The Telegraph. https://www.telegraph.co.uk/global-health/terror-and-se

curity/live-streamed-child-abuse-philippines-surges-britain-demand/

2. New York Times. (2024). On these apps the dark promise of mothers sexually abusing children. https://www.nytimes.com/2024/12/07/us/child-abuse-apple-google-apps.html

3. International Justice Mission & University of Nottingham Rights Lab. (2023). Scale of Harm Research Method, Findings, and Recommendations. https://ijmstoragelive.blob.core.windows.net/ijmna/documents/studies/IJM_Scale_of_Harm_2023_Full_Report_5f292593a9.pdf

4. Europol. (2024). Internet Organised Crime Threat Assessment (IOCTA) 2024. https://www.europol.europa.eu/cms/sites/default/files/documents/Internet%20Organised%20Crime%20Threat%20Assessment%20IOCTA%202024.pdf
5. Childlight. (2023). Searchlight 2023: Flagship Report. https://www.childlight.org/uploads/publications/Childlight-Flagship-Report-2023_1.pdf

6. Australian Institute of Criminology. (2023). The Overlap Between Child Sexual Abuse Live Streaming, Contact Abuse, and Other Forms of Child Exploitation. https://www.aic.gov.au/sites/default/files/2023-05/ti671_overlap_between_csa_live_streaming_contact_abuse_and_other_child_exploitation.pdf

7. First Global Ministerial Conference on Ending Violence Against Children. (2024). Government Pledge – United Kingdom. https://endviolenceagainstchildrenconference.org/wp-content/uploads/2024/11/United-Kingdom-pledge.pdf

8. Online Safety Act. (2023). UK Parliament. https://www.legislation.gov.uk/ukpga/2023/50/contents/enacted