LUCY FAITHFULL FOUNDATION

Funded by Nominet

TRIGGER FOR CHANGE

The power of warnings to prevent online child sexual abuse

Project Intercept Impact Report

Sarah Smith, Frances Frost, Tom Fisher

> “In a world where there’s so much negative news about child sexual abuse material, it’s so refreshing and reassuring as a survivor to hear about the great work being done in Project Intercept.”
>
> Member of the Lived Experience Group, Marie Collins Foundation

> “There is a misconception that encrypted services cannot address harm. In reality, we take meaningful action, including delivering deterrence messaging via Project Intercept, a science-backed approach we value.”
>
> André Meister, Chief Technology Officer (NZ), Mega

> “Seeing the warning message made me call the helpline.”
>
> Anonymous caller to **Stop It Now**

LUCY FAITHFULL FOUNDATION

Contents

4 **Executive summary**

6 **What is Project Intercept?**

8 **Our impact**

8 Disrupting harm with warning messages

18 Encouraging help-seeking behaviour

28 Evaluating Intercept’s impact on tech companies

32 Influencing regulation

34 Enhancing research into online child sexual abuse

36 **Looking ahead**

37 **Partner with us**

Funded by Nominet

Acknowledgements

Thank you to **Nominet** for funding Project Intercept, sharing our vision and making the impact showcased in this report possible.

We are also grateful for the support and strategic guidance from **Project Intercept's Consortium Board: University of Tasmania (UTAS)**, **International Policing** and **Public Protection Research Institute (IPPPRI)**, **Internet Watch Foundation (IWF)** and law enforcement partners. In particular, thank you to Joel Scanlan who has gone above and beyond in his support of Project Intercept.

A heartfelt thank you to **Thorn** and the **Tech Coalition**, who quickly championed the work and opened the doors to partnerships that made the work happen.

Partners

Thank you to the tech partners shown here, who have collaborated with us and opted to share their participation publicly. We also engage with additional partners who are not listed here but deliver the same interventions.

Executive Summary

Over the past two years (2024 and 2025), **Project Intercept** has shown that warning messages can play a key role in preventing online child sexual abuse. By interrupting harmful behaviour at the moment of risk, the project has diverted hundreds of thousands of people away from offending and towards confidential support at our **Stop It Now** helpline and online self-help.

Developed by the Lucy Faithfull Foundation (LFF) in response to the rapid growth of child sexual abuse material (CSAM) online, **Project Intercept** focuses on prevention rather than reaction, working with technology companies to challenge behaviour before abuse occurs.

Project Intercept's vision is bold: what if every time someone tried to access harmful content on the internet, they encountered a message designed to stop them and encourage them to seek help at **Stop It Now**?

Two years on:

- Our messages have been viewed **more than 70 million times globally.**
- **Nearly 700,000 people have clicked through from our warnings to seek support.**
- Each month, **an average of 28,000 users engaged** with our self-help resources, with an **81% average engagement rate**.
- **Our messages have reached 131 countries.**
- We've moved far beyond simple search interventions and even operate successfully in **encrypted** environments.
- We own and operate **two deterrence chatbots**, one in English and one in Dutch.
- We are conducting **new research into user pathways** to help us improve warning messages and add learning for the sector.
- Our evidence has helped strengthen recommendations in **Ofcom's illegal harms codes**.

From global search engines to emerging AI platforms, online tech companies are proving that warnings can be deployed at scale. But there's more to do. If you're a tech company and not yet working with us, now is the time. Warning messages are low-cost, high-impact, and can be tailored to your platform. Partner with us to strengthen your trust and safety framework and protect children worldwide.

Read on to see the evidence and the impact **– then join us**.

What is PROJECT INTERCEPT?

The scale of online child sexual abuse is vast and persistent. In 2025 alone, the Internet Watch Foundation took action on more than 312,000 webpages containing CSAM – each page potentially hosting multiple images or videos of abuse. While detection and removal are essential, the volume of material makes clear that reactive measures cannot keep pace on their own.

LFF was established in 1992 to prevent child sexual abuse in all its forms. We combine research, clinical expertise and frontline delivery, working with people who pose a risk of harm, those who have offended, and the professionals responsible for safeguarding children. Through our confidential Stop It Now services and targeted prevention campaigns, we reach thousands of people each year at risk of causing harm.

Project Intercept is our response to the online dimension of this problem. It is a practical deterrence initiative designed to interrupt harmful behaviour. Built on original research into the effectiveness of targeted messaging, Project Intercept works directly with technology companies. These include some of the world's largest platforms, as well as services operating in more complex environments such as end-to-end encrypted spaces.

Why Project Intercept and why now?

For years, technology companies have focused – rightly - on detecting and removing illegal content. But far less attention has been paid to stopping the behaviour that creates that content. When Project Intercept began, warnings, where they existed, were often reactive, deployed after harm had already occurred, or aimed primarily at potential victims.

We asked different questions:

What if you put friction in the pathway of the offender?

What if you disrupt the behaviour, make the consequences clear, and offer a route to help?

Making that happen was not straightforward. Embedding help-oriented deterrence inside complex digital products required trust, collaboration and shared learning. At the outset, it was far from certain that technology companies would engage.

Project Intercept exists because tech companies *did* engage - trust and safety teams were prepared to examine uncomfortable gaps, pilot interventions inside live environments, and measure the results. What began as a question (*can we interrupt harmful behaviour at the point of risk?*) has resulted in a growing body of evidence, partnerships and real-world implementation.

The funding approach that helped Project Intercept thrive

Project Intercept has been able to grow in ways we could never have predicted, and a large part of that is down to the funding that has supported it. Instead of tying us to rigid targets, our funder, Nominet, backed our vision. Their Countering Online Harms Innovation Fund enabled us to test and learn about what works in preventing online sexual harm to children and adapt quickly when the online landscape shifted. None of this would have been possible under a more rigid funding model.

The funding also strengthened innovation at LFF. Support from CAST[1] (Centre for the Acceleration of Social Technology), along with new tools and ways of working developed through our Innovation Hub, gave us the confidence to explore emerging technologies and expand our thinking.

Key things that made this funding so effective

A focus on continuous testing, learning and adapting rather than rigid KPIs

Funding paid up front, reducing pressure and enabling early hiring

Wrap-around support from CAST as our innovation mentors

Funded by Nominet

[1] https://www.wearecast.org.uk/about-us/

OUR IMPACT

DISRUPTING HARM with warning messages

AT A GLANCE:

70 MILLION
triggers of **Project Intercept** warnings

22
active warnings

11
different sectors of the tech industry

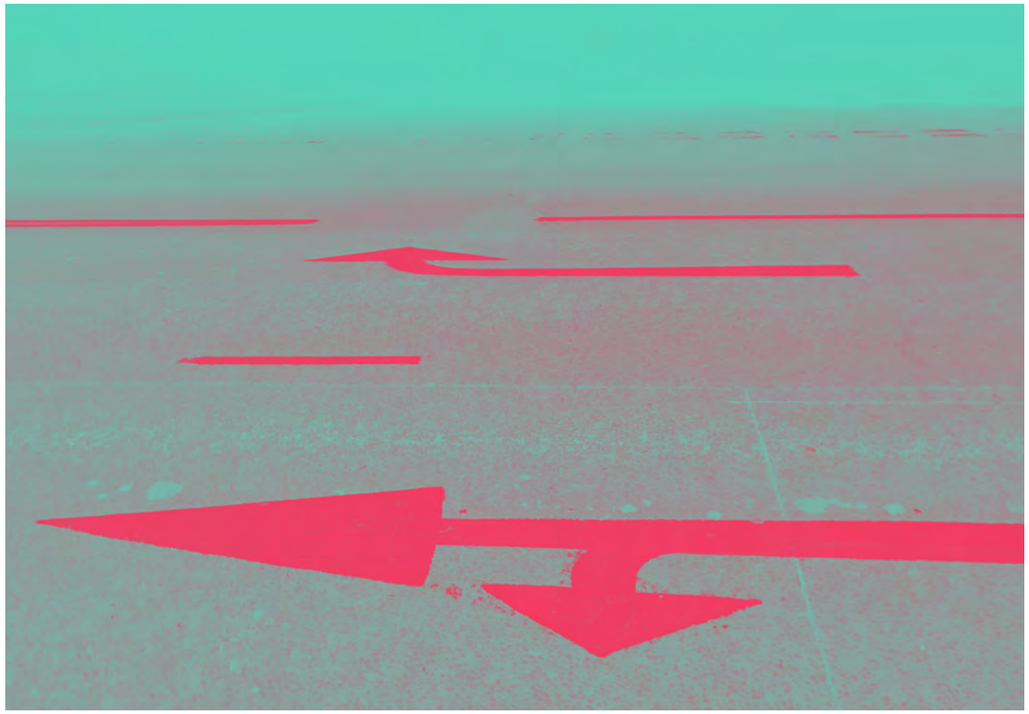
Tech companies are often reluctant to publish data on how and when they deploy warnings. However, based on information shared by Project Intercept partners, we estimate that our warnings have been deployed at least 70 million times over two years.

This figure is rather sobering when viewed through a victim's lens. Each one of those 70 million warnings represents someone attempting to access illegal material that causes harm to children.

Yet, viewed through a preventative lens, it is a significant success story – 70 million attempts to cause harm have been intercepted, and each time an offer of support to change behaviour has been given. Where that offer has been taken up, there has been strong engagement with our online self-help resources.[2]

[2] 81% of people stayed longer on the resources than 10 seconds, and took at least one action (eg clicking, scrolling, opening modules)

How warning messages work



A wide reach

Over 2024 and 2025 we engaged 50 tech companies in 11 different sectors. To date, this engagement has resulted in 22 active warning interventions, with some partners hosting more than one.

Some warnings operate only in the UK, while others have a global reach, deployed wherever a platform is active. The list to the right includes sectors where warnings are already live, as well as those where interventions are currently in development.

- Search
- Artificial Intelligence
- Cloud storage
- Gaming
- Social media
- Adult
- Dating
- Internet Service Providers
- Music streaming
- Video platforms
- Messaging

Example page from the **Get Help** programme

A note on the Get Help programme

When users click one of our warnings they are redirected to the opening page of our free, confidential self-help programme, **Get Help**, housed on the **Stop It Now** website.

The programme, which has been running for 12 years, comprises an extensive range of psycho-educational modules designed to help a user confront, reflect and start to change their behaviour. The modules are designed to be revisited so the user can work at their own pace and can be used in conjunction with support from our trained advisors over email, live chat or via the **Stop It Now** helpline.

This quote from one of our service users, Ben[3], shows clearly the relationship between warnings, our online self-help and its efficacy.

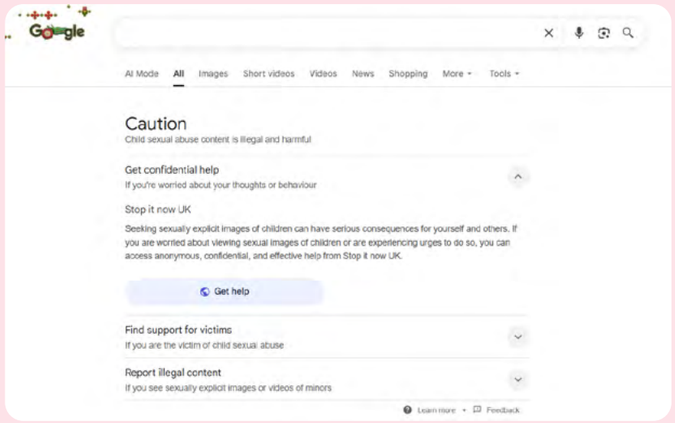
“Over the years I have been in and out of depression and recently suicidal. I used adult sites as a way to pass time and quiet the mind. I had a pop-up [warning] on an adult website… I had a look around at what you did and read some of the modules.

About two months ago I gave up those sites and decided I want to keep my mind occupied and more productive. I found the modules on addiction and pornography very helpful. And since I am free from it I feel better in myself.”

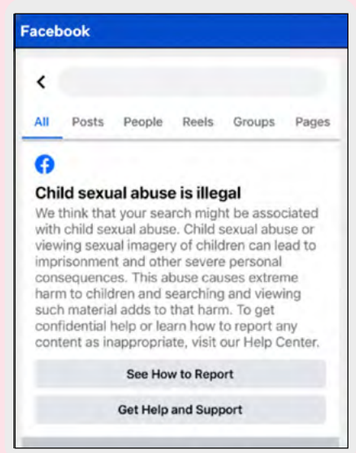
[3] Ben is a pseudonym

Examples of our WARNINGS

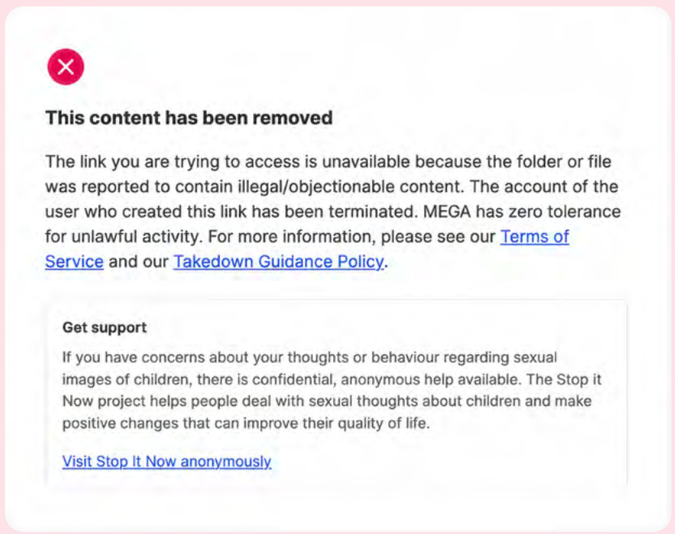
Most partners use their own branding but the overwhelming majority take advice from our clinical team and our research partners at UTAS on the wording and tone used. See some examples below:



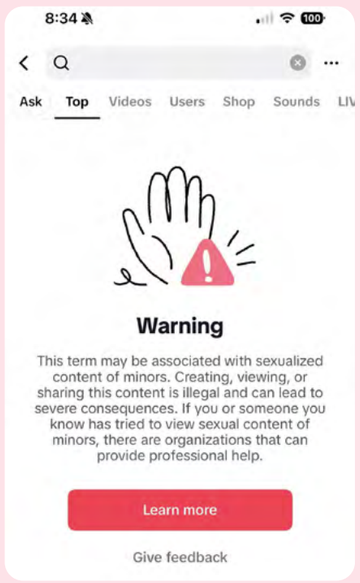
Google warning



Facebook warning



Mega warning



TikTok warning

What makes a good warning message?

Before we began our first deterrence campaign in 2015, LFF carried out research with arrested offenders to understand, from their point of view, the essential components of creating an effective warning message.[4]

From their responses we found that the key messages to include were:

It's a crime

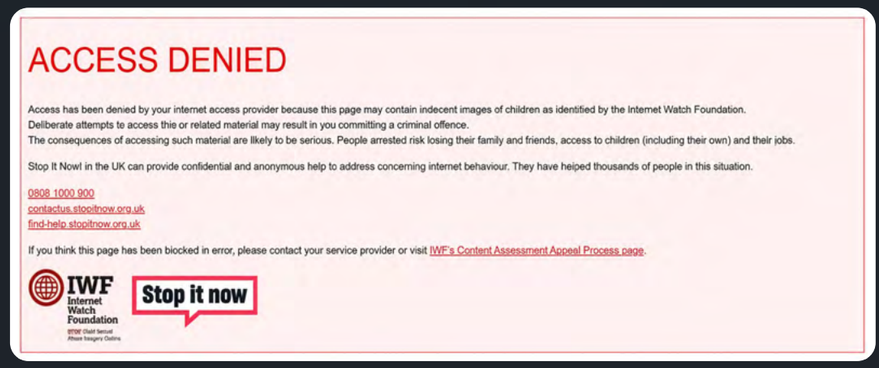
It causes harm to children

There are consequences

But there is help available

We incorporate these ideas into each message.

[4] Bailey, A., Allen, L., Stevens, E., Dervley, R., Findlater, D., & Wefers, S. (2022). Pathways and prevention for indecent images of children offending: A qualitative study. Sexual Offending: Theory, Research, and Prevention, 17, 1-24. https://doi.org/10.5964/sotrap.6657



Typical splash page served by an ISP when a user attempts to access a webpage that has been blocked for hosting CSAM

The LFF Chatbot – a deeper layer of deterrence

Deterrence messaging isn't just about static warnings. In 2024 we took on ownership of the **reThink Chatbot**, a tool originally developed in partnership with the IWF and Aylo (the owners of Pornhub).[5]

The chatbot still sits on Pornhub in the UK and seeks to engage users searching for CSAM. In 2025, we developed a Dutch version of the chatbot in partnership with Stop It Now! Flanders, which is hosted on Pornhub in Belgium. We are also developing a generative AI version of the chatbot trained on motivational interviewing techniques, in partnership with the Karolinska Institutet and Orebro University in Sweden.

One legacy of the original reThink chatbot is the major findings from its 2024 evaluation[6], which give insights as to the deterrence impact of the chatbot, not just the impact on help-seeking behaviour:

- the presence of the chatbot and warning message helped to drive down the number of searches for CSAM over the 18-month evaluation period.
- 82% of sessions needed only a single encounter with the bot and warning message for users to stop their offending behaviour in that moment.

The deterrence impact

Although the evidence from our chatbot comes from the adult sector, we anticipate that similarly strong deterrent effects could occur across a variety of online spaces when warnings are deployed.

Further research in collaboration with tech companies is needed to grow the evidence base, but the evaluation of our chatbot indicates that over time warnings will drive down the volume of attempts to access illegal material. Additionally, most users who encounter a warning will stop their behaviour in that moment.

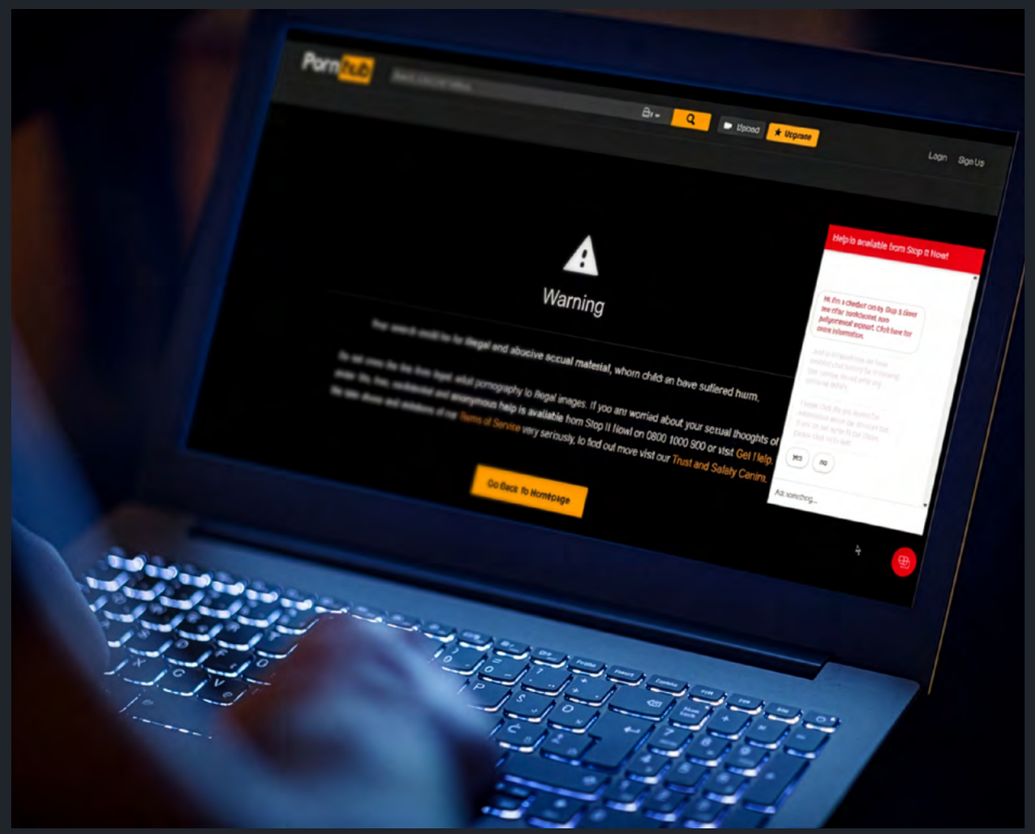
[5] https://www.lucyfaithfull.org.uk/research/chatbots-and-warning-messages-innovations-in-the-fight-against-online-child-sexual-abuse/

[6] Scanlan, J., Hall, L. C., Watters, P. A., Wortley, R., & Prichard, J. (2026). The Relationship Between a CSAM Warning Messaging Chatbot and User Behavior on Pornhub. Victims & Offenders, 21(3), 574–602. https://doi.org/10.1080/15564886.2025.2606857

Opportunity to scale

We are incredibly proud that through Project Intercept, tens of millions of users have been served warnings from just 22 interventions. But the potential to scale is huge. With more warnings in more online spaces, millions more could be reached, achieving mass-scale disruption.

Pornhub warning



This is what **70 MILLION WARNINGS** are for

One real case.
One moment of intervention.
One man who chose to change.

Simon, a 25-year-old man from Peru, was a regular CSAM user. He first saw pornography at the age of seven, and was drawn to CSAM use in his early 20s. One evening, he attempts to access CSAM on the cloud storage site Mega.

Mega, a Project Intercept partner based in Hungary, has already detected and disabled the CSAM. But Simon's attempt to access the files triggers a warning.

The warning tells him his search is for illegal material, that it harms children and that there will be consequences to his behaviour. It also suggests he seek help at **Stop It Now**, if he is ready for change.

Simon has been regularly struggling with guilt and shame, and has deleted his collection of CSAM several times. Somehow, though, he is always drawn back to the behaviour.

He clicks the link on the warning and starts reading the information on Stop It Now. Using the built-in translation tool he can read the whole site in Spanish.

He then starts a live chat with our advisor based in the UK, using a translation tool as he types.

He tells our advisor he has no friends and is living alone, using CSAM regularly. He wants to change but he is isolated and doesn't know how.

The advisor reassures him that change is not only possible but crucial. She guides him through a set of suggested actions to take, including using our online self-help programme, **Get Help**. She encourages him to chat or contact us regularly for support and accountability.

Feeling supported, Simon is motivated to change his behaviour and begins a journey to better choices.

Encouraging HELP-SEEKING behaviour

AT A GLANCE:

ALMOST
700,000
people clicked through for help

81%
Engagement rate

131
Countries reached

Over two years, 2024 and 2025, we have seen 682,730 users click through to our self-help resources from warning messages deployed by our partners. Through tracking links embedded in the warnings themselves, we know that approximately 28,000 users each month are diverted and come through to our Stop It Now website when a warning has been deployed.

Users are engaged when they arrive. Coming from 131 countries, those users have together amassed more than two million page views with an average engagement rate of 81%[7]. This high engagement rate means that when arriving at our Stop It Now website[8], most users are not leaving the website right away but are having meaningful interactions with the resources available, including our online self-help modules.

A further hard-to-measure impact relates to those who do reach out for help but not via the warning itself. There is likely to be a cohort who mistrust the offer of help within the warning and instead choose to independently access Stop It Now, perhaps in the minutes, hours or even days after viewing a warning. As this cohort seeks help without engaging our tracking links, it is difficult to estimate its size and nature.

It is important to note that help-seeking behaviour is only ever a small part of the impact that a warning has. The vast majority of those served a warning will not click through for help. Most are simply disrupted – the deterrent effect. Therefore we believe that the almost 700,000 users that have clicked through for help in the past two years from Project Intercept warnings represent one per cent of a much larger number, 70 million, whose behaviour online was confronted when they were served a warning.

[7] Engagement rate is the percentage of website sessions that are considered engaged. An engaged session lasts longer than ten seconds and includes one or more page views and/or at least one conversion. It is calculated by dividing engaged sessions by total sessions.

[8] An independent evaluation of the Stop It Now helpline found clear demand for the service and evidence that it provides quality, cost-effective advice that can prompt behaviour change and strengthen protective factors that reduce the risk of harm. NatCen Social Research (2014), A public health approach to tackling child sexual abuse: Research on Stop It Now UK and Ireland, funded by the Daphne III programme of the European Union. A further independent evaluation of the Stop It Now helpline is currently underway.

Emergent themes from our data

Triggering is key
Getting the triggering right (the rules that determine when a warning is activated) for deploying a warning is particularly important to its effectiveness.

Triggering that is too broad will activate warnings for behaviour that is not harmful, while triggering that is too specific will fail to identify some people attempting to access illegal content.

For search-based interventions, maintaining and updating a keyword list is critical. Through Project Intercept, we have been able to direct our partners to NGOs such as the IWF and Thorn, who are best placed to support this.

Decreased engagement over time
The evaluation of our chatbot showed that although the volume of searches that triggered the intervention decreased over time, so did the requests for information about our Stop It Now services, at an even faster rate. This could indicate that over time, as users become accustomed to seeing warnings, they start to take less notice of them.

This also holds true for our most successful warning in Project Intercept in partnership with Mega, where the percentage of those clicking through on the warning has decreased over time.

Optimisation
Although engagement with warnings may decrease over time, this presents an opportunity. Refreshed warnings have demonstrated renewed effectiveness, successfully reversing habituation.

This has been highlighted through our longest-standing warnings in search, with Google and Bing, where multiple versions of the warnings have been implemented over the years. This work shows that by introducing thoughtful changes, it is possible to improve the effectiveness of warnings in driving help-seeking behaviour.

Deterrence at scale
INSIDE ENCRYPTION

MEGA

Our most prolific warning runs with the cloud storage company Mega. This partnership is significant for two reasons. First, law enforcement had flagged Mega as a platform frequently used by CSAM offenders. Second, Mega operates a fully end-to-end encrypted service, long seen as a space where meaningful intervention is not possible without compromising user privacy. Our work together has shown this assumption is wrong.

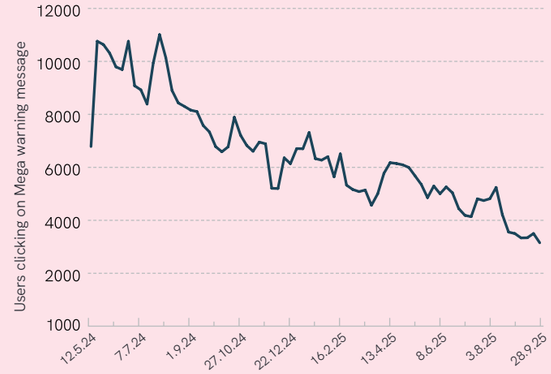
High visibility, immediate impact
We connected with Mega in spring 2024 via the Tech Coalition. The warning went live on 9 May 2024 and the response was immediate. Thousands of users clicked through to Stop It Now in the first days, with high engagement on our website. Since launch, Mega has become the backbone of our deterrence data. The warning's clickthrough rate consistently sits between 1 and 2 per cent, meaning millions of users see the message each year.

What Mega does differently
Mega's approach is unique. Because the platform cannot see user content, it relies on public reports of CSAM links. When a report is made, the hosting account is disabled, and anyone attempting to access the link is served a warning. No content is viewed; no privacy is compromised.

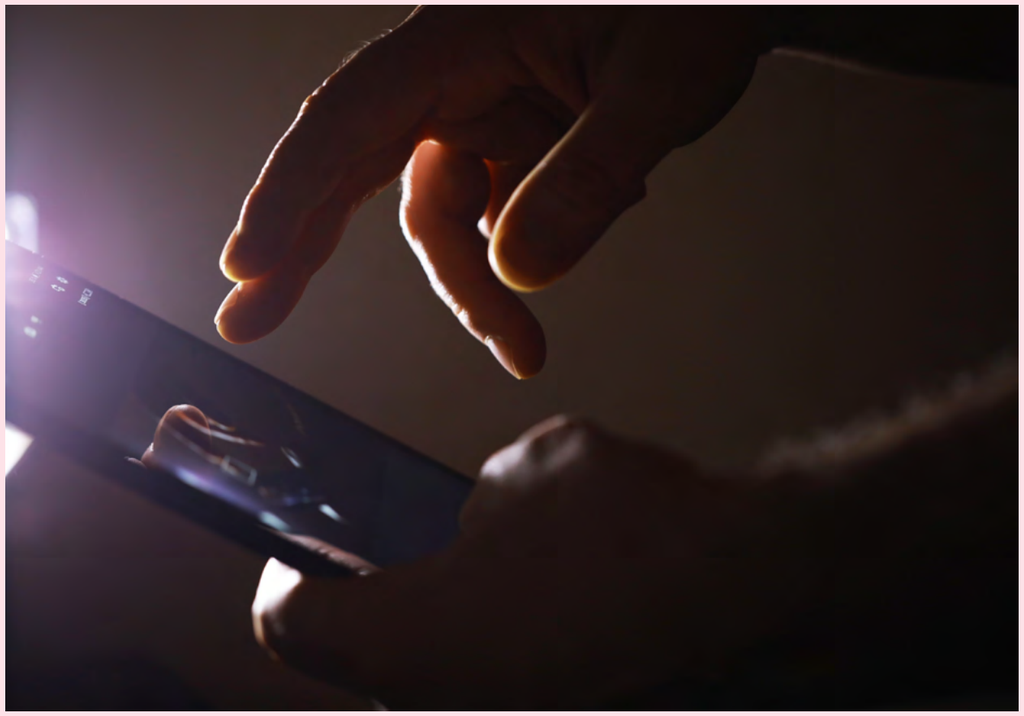
Users suspected or confirmed to be storing CSAM are also sent emails reinforcing the warning message and offering a route to help. Appeals that are rejected trigger similar messages.

Impact over time
Over time, clickthrough numbers to Stop It Now have gradually reduced, reflecting Mega's success with various initiatives, including deterrence, in reducing the spread of objectionable content on its service. However, the clickthrough rate has also declined over the period since the intervention went live. This may indicate habituation – the more times the warning is viewed, the less effective it is in engaging users and directing them to seek out the support offered.

This case study demonstrates that effective deterrence is possible even within encrypted environments, offering a practical model for other platforms to follow.



Graph showing gradual decline in clickthroughs to Stop It Now from Mega warning

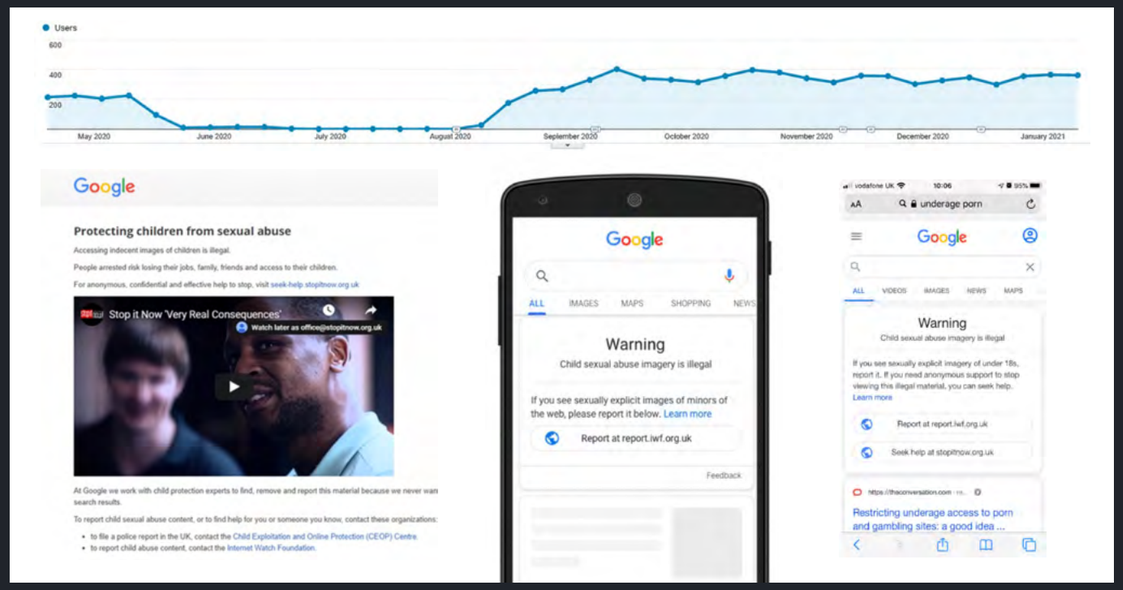


How a NATURAL EXPERIMENT shaped the future of warnings

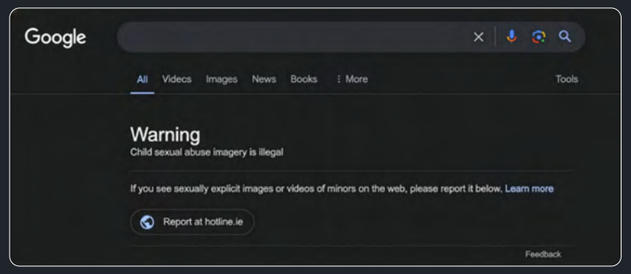
Google introduced one of the earliest deterrence warnings in UK search in 2013, signposting people searching for harmful material to support at Stop It Now. By 2018, Google were reporting a thirteenfold reduction in CSAM-related search queries[9].

During the Covid 19 pandemic, click-throughs to Stop It Now suddenly dropped from around 200 per week to zero. Google later confirmed that a design change, intended to simplify the warning through clearer wording, had accidentally removed the link to Stop It Now. When the link was reinstated in autumn 2020, traffic immediately returned and, at times, doubled. This natural experiment showed how design and language directly affect user behaviour and was a source of inspiration for Project Intercept.

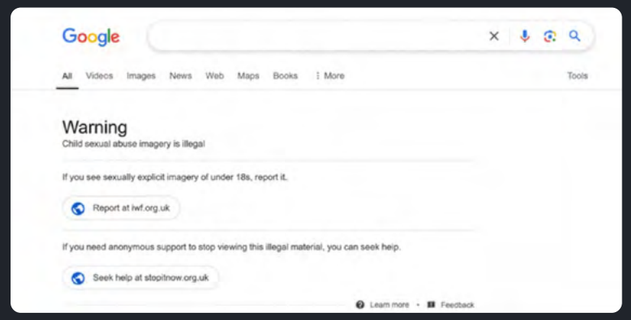
[9] https://blog.google/innovation-and-ai/technology/safety-security/continuing-fight-against-child-sexual-abuse-online/



The effect of re-designs
Further evidence emerged in 2024, when a new Google warning implemented in Ireland generated very few clickthroughs to Stop It Now. The key difference was design. In Ireland, support was hidden behind multiple layers. The UK warning, which was performing much better, still placed the Stop It Now link beneath a prompt to report CSAM to the IWF, a structure we felt risked obscuring the route to help.



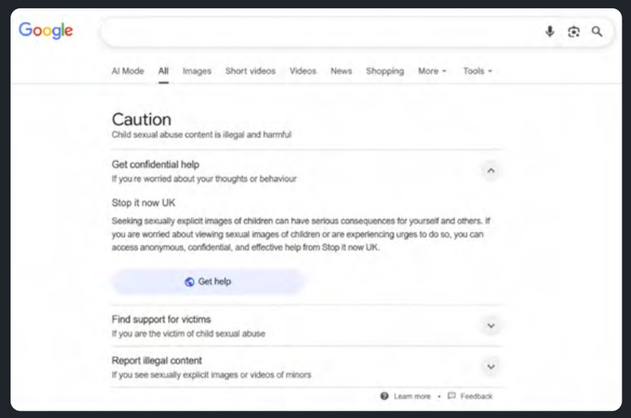
Google Ireland warning, prior to February 2025



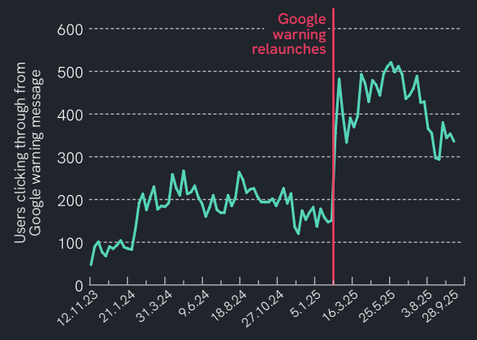
Google UK warning, prior to February 2025

In response, Google unified its warnings globally in February 2025, placing help-seeking at the top and using more engaging language. Traffic increased in both the UK and Ireland.

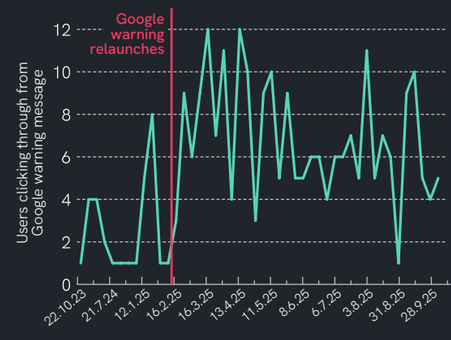
This case study shows how small changes to wording, placement and the visibility of help can dramatically affect whether people choose to seek support.



Google UK warning traffic



Google Ireland warning traffic

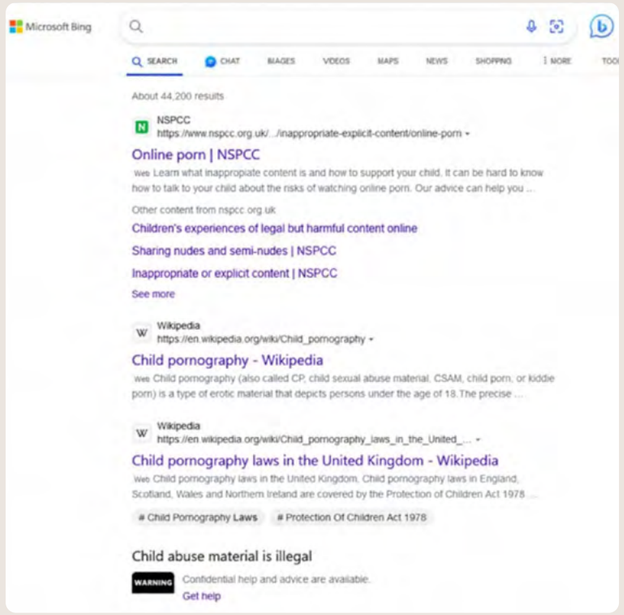


REFINING THE MESSAGE:
Lessons from our work with

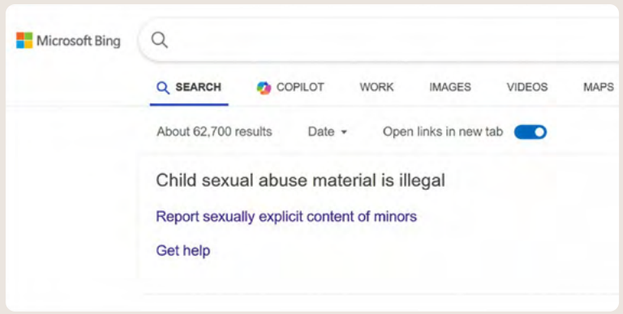
Microsoft Bing

Our partnership with Bing began in 2013 but became inactive over time. It was revived in 2024 as part of Project Intercept's work to optimise existing warnings.

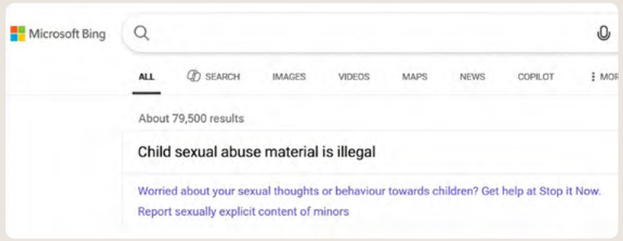
Early conversations showed the original warning was no longer effective. It sat at the bottom of the search results page, was easy to miss, and lacked a tracking link, making its impact hard to assess.



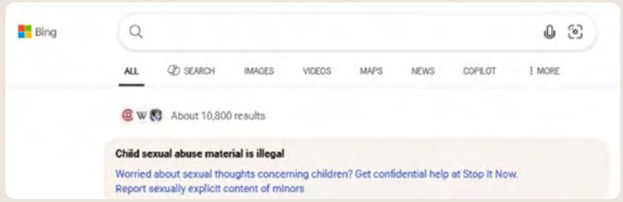
As Bing updated warnings across multiple countries, we worked with their team to refine language and move the warning to the top of the page, clearly separated from search results, so that users could find authoritative reporting and support options at one glance. A new tracking link was added, enabling ongoing monitoring.



In summer 2025, we advised on clarifying the call to action by explicitly signposting to Stop It Now. This change led to an immediate rise in the number of clickthroughs to Stop It Now.



In autumn 2025, we recommended including the word "confidential" and supported an international rollout of the warning, tracked separately. Our multilingual self-help modules, hosted on the Stop It Now website, guide users through psycho-educational exercises to support behaviour change. Because they are available in more than 40 languages, they offer a scalable solution that technology companies can deploy globally across regions where no local service exists, and they are generating regular traffic.



After these changes were implemented, we saw an overall 47% increase in the UK clickthrough rate in Bing search with good engagement on the Stop It Now site. This work shows how small, evidence-led changes can significantly improve the impact of deterrence messages.

Evaluating Intercept's impact WITHIN TECH COMPANIES

AT A GLANCE:

- Trust & safety professionals report a step change in how they understand child sexual abuse prevention
- Trust & safety professionals are better equipped to influence senior decision-makers and legal teams
- Help-oriented prevention is becoming more normalised inside tech companies

How we gathered these insights
This chapter draws on in-depth conversations with 14 trust and safety professionals working across a range of technology companies. The interviews, carried out by the LFF research department, form the basis of our forthcoming Project Intercept evaluation report.

Some interviewees were active Project Intercept partners. Others had explored working with Project Intercept but were unable to progress, often because of organisational constraints such as resourcing or competing priorities.

Shifting how prevention is understood
For many professionals, engaging with Project Intercept marked a turning point in how they thought about preventing child sexual abuse. Several spoke openly about starting from a place of uncertainty or scepticism, and then reassessing that position as they learned more about how help-oriented deterrence works.

> **"Working with Project Intercept has brought me around to the side of prevention more… I don't really think I was much of a believer in it, if I'm being honest. I don't think I was really convinced of its effectiveness, and I think that the work with Lucy Faithfull made me much more open to it and really supportive of what they were doing."**
>
> Professional from a social media platform

This shift often meant moving beyond a narrow focus on enforcement or content removal, towards a clearer understanding of how online offending develops and where interventions can make the greatest difference.

Professionals were also candid about the fact that this way of thinking was not always intuitive, particularly in trust and safety environments shaped by enforcement-led practice.

> **"Help-oriented deterrence is kind of a confusing concept at first. You don't really know how to grapple with it… It was a case of, 'why are we doing anything to help these people?'"**
>
> Trust and safety professional at a technology company

What stands out is not that these questions existed, but that engagement with Project Intercept helped professionals work through them rather than stopping at doubt.

Building confidence and credibility
As understanding deepened, professionals described feeling better equipped to make the case for prevention internally. Project Intercept's impact here was practical as well as conceptual, helping people frame prevention in ways that resonated with organisational priorities.

> **'It's helped me be much more constructive in arguments around resourcing, because I'm able to articulate the returns that can be made on reducing overall prevalence of harm, and in terms of value to our brand as well…'**
>
> Former Trust and Safety Lead at a social networking company

A critical part of this confidence came from data. Professionals repeatedly pointed to the value of real-world behavioural evidence from live warning messages. Seeing how users respond in practice - not just in theory - helped overcome internal scepticism and reduced the perceived risk of trying something new. For decision-makers weighing legal, reputational and resource concerns, evidence from data made prevention feel measurable, defensible and worth investing in.

Being able to ground conversations in evidence changed their tone. Prevention was no longer just the right thing to do; it was something that could be shown to work.

From confidence to action
In some cases, working with Project Intercept supported the introduction of new CSAM warning messages, establishing a new internal precedent. In others, the impact lay in optimising existing interventions, refining wording, tone or placement to strengthen their preventative effect.

> **Because it came from the Lucy Faithfull Foundation and the expertise they have in this field… and there was research behind it… that made us more at ease in making these changes.**
>
> Search engine professional

This was not just about implementation. Professionals described feeling more comfortable navigating internal processes, particularly where legal or reputational concerns were raised. Even where change could not happen immediately, Project Intercept helped keep prevention visible and professionals felt better prepared when opportunities did arise.

> **Working with Project Intercept was a great experience. They're very thoughtful with the data. They want to make sure we understand… so they really contextualise.**
>
> Trust and safety professional

Normalising prevention within trust and safety
Taken together, these experiences could point to the start of a gradual but meaningful shift in trust and safety culture. This would mean that help-oriented prevention is becoming easier to discuss, easier to justify and more widely accepted as part of everyday practice.

A two-way effect beyond CSAM
Project Intercept's influence also connects with wider safety cultures inside technology companies. In some organisations, CSAM warnings were the first help-oriented deterrence messages introduced, and their success inspired similar approaches in areas such as suicide or self-harm.

In others, prior experience with deterrence in those areas created a foundation that made it easier for Project Intercept to introduce CSAM warnings. In both cases, child sexual abuse prevention became part of a broader conversation about early intervention and safety by design.

Laying the groundwork for wider change
Cultural change in large technology companies is rarely quick or straightforward. What is visible here is steady progress: growing confidence, stronger internal advocacy and a clearer sense that early, help-oriented prevention belongs within trust and safety practice.

Together, we have seen these shifts start to lay the groundwork for longer-term change across the sector.

Influencing REGULATION

AT A GLANCE:

Ofcom strengthened its guidance on warning messages following evidence from **Project Intercept**.

Project Intercept has contributed evidence at a key moment in the development of online safety regulation. The UK's Online Safety Act became law in October 2023, and Ofcom published its consultation on the Illegal Harms Codes in November 2023. Evidence from Project Intercept was submitted as part of our response to this consultation, bringing real-world evidence into regulatory discussions about prevention.

This evidence directly informed Ofcom's final position. While the consultation document referenced the Lucy Faithfull Foundation five times, the published Illegal Harms Statement in December 2024 referenced our work 37 times, reflecting the weight given to evidence from Project Intercept. Ofcom also strengthened its guidance to state that technology companies should consult experts when designing warning messages, recognising the role of specialist insight in effective prevention.

> **"We consider it reasonable to expect that a warning message developed in collaboration with experts (as suggested by the Lucy Faithfull Foundation) would be more effective at deterring searches for CSAM and providing appropriate support than a message developed without this collaboration. We have decided to amend the measure to recommend that service providers develop their warning messages with input from a person with expertise in deterring CSEA[11] offences."[12]**

Only large search services are currently required to use warning messages and signpost to support. Evidence from Project Intercept shows that messaging can be effective across a much wider range of services and sectors, and we continue to share this evidence as regulation evolves.

> **"We welcome innovation… in this space and will keep this area under review for future iterations of our Codes as the evidence base grows."[13]**

Across the world there is a wider shift towards stronger online safety regulation. Regulators are increasingly focused on platform responsibility, safety by design, and early intervention, rather than relying solely on detection and removal after harm has occurred. Evidence from Project Intercept continues to contribute to this broader conversation.

[11] Child Sexual Exploitation and Abuse.

[12] Ofcom (2024). Protecting people from illegal harms online: Illegal Harms Statement, Vol. 2: Service design and user choice, para. 9.99.

[13] Ofcom (2024). Protecting people from illegal harms online: Illegal Harms Statement, Vol. 2: Service design and user choice, para. 9.143.

Enhancing RESEARCH into online child sexual abuse

AT A GLANCE:

- In-depth interviews with **14 prisoners and arrested offenders**, providing rare insight into behaviour, motivation and decision-making
- A survey of more than **3,000 CSAM forum users** on the dark web, strengthening the evidence around online offending patterns
- AI-supported analysis of **dark web forum discussions**, identifying common themes
- Behavioural analysis of pornography users who search for CSAM, informing how and when **warning messages** can be most effective
- A **pilot deterrence survey** on warning messages undertaken with users of LFF Programmes

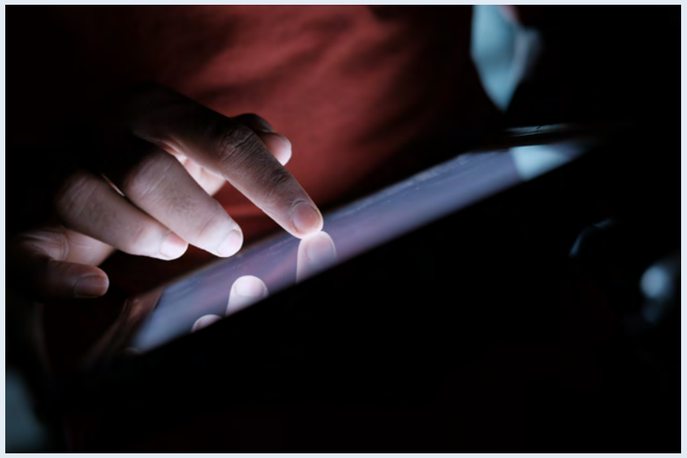
Research has been central to Project Intercept from the start. Two major research programmes have run in parallel, together addressing some of the most significant evidence gaps in the field of online child sexual abuse prevention.

The **Prisoner Debriefing Study** involved researchers visiting prisons across the UK to interview people convicted of CSAM and online grooming offences. These interviews explored how individuals first encountered CSAM, what drew them further in, and how they navigated online spaces. This under-researched group offers critical insight into how warning messages should be designed, where they should appear, and the tone most likely to disrupt harmful behaviour.

Alongside this, a large-scale research programme, **Pathways to Offending**, has been delivered for Project Intercept by the International Policing and Public Protection Research Institute (IPPPRI) at Anglia Ruskin University. There are three strands:

- Analysis of behavioural data from a major pornography site, examining patterns of activity before and after users encounter a warning message.
- Web scraping and thematic analysis of discussions across multiple dark web CSAM forums, identifying recurring narratives and rationalisations.
- A survey conducted within a dark web CSAM forum, exploring users' motivations and attitudes towards risk and intervention.

Results from both studies are due to be published later in 2026. The findings will help Project Intercept focus its remaining work in the places where it can have the greatest effect. It will also provide evidence to support other organisations working in the field.



Looking AHEAD

AT A GLANCE:

- **Experimenting** with different types of warning messages in the adult and AI sectors
- Creating a **generative AI chatbot** powered by motivational interviewing techniques
- Bringing more **tech platforms** on board to scale our work
- Publishing our **original research** on warnings
- Continuing to influence the **regulator** to mandate wider use of warnings

There are now 13 months left in Project Intercept, which is due to finish in June 2027. There's still much to achieve.

As well as continuing to bring in new tech partners and scale up our warnings, we're launching some in-depth **experimentation** with two of our most trusted tech partners. Using A/B testing[14] we will experiment with different approaches to language and design of messages that will bring us closer to understanding what works best in specific online spaces. This work will give us robust behavioural evidence, showing how users respond in real time, not just how they report their intentions in online surveys.

We will publish the results of our research and use the findings to influence how we develop messages in the final year of our work.

And we will deliver our Motivational Interviewing chatbot, which has the potential to create a whole new body of evidence for online interventions.

As ever, we will continue to influence how platforms think and act regarding prevention, pushing for clearer, more consistent approaches to child protection in both tech practice and regulation, so that the online world becomes a safer space for children.

[14] A method of comparing two versions of something to determine which performs better. People are randomly shown one version or the other, and the results are measured to identify which achieves stronger outcomes.

Partner with us

Intercept would be nothing without the technology platforms who have supported our work. We are always open to hearing from new potential partners. If you have questions or just want to see how we might be able to help, please get in touch. See our FAQ for more guidance.

LUCY FAITHFULL FOUNDATION

FAQ

How long does it take to get a partnership off the ground?
We work to our partners' timelines and are guided by you. Some of our warnings have been conceived, designed and are operational within a matter of weeks. Some have taken more than a year.

What sectors do you work in?
We will work within any online services or platforms that are experiencing harms. We know that harm is happening across the whole of the online ecosystem. Wherever there is a risk of harm being caused, we will try to find a way to create interventions.

What does partnership look like?
There's no set process, but we typically start by understanding how your platform operates and where the risks are. We review what you're already doing and suggest improvements, such as refining existing messages. We then develop tailored warning content using our expert partners and clinical team. Once live, we report back regularly with data and insights on effectiveness.

Do we have to be open about our work?
All our work with our partners is confidential. Those named in this report have agreed to be included and approved the use of their logos. While some organisations choose to publicise their work to prevent harm, we understand this is not right for everyone, and we have several partners who have chosen to remain private. How visible our partnership is will be led by you.

How much does it cost?
Our work is currently free of charge, as we are kindly funded by Nominet. The only cost to you is in staff engineering time.

Ready to join us?
Contact: **intercept@lucyfaithfull.org.uk**

LUCY
FAITHFULL
FOUNDATION

The Lucy Faithfull Foundation is a Registered Charity No. 1013025, and is a company limited by guarantee, Registered in England No. 2729957