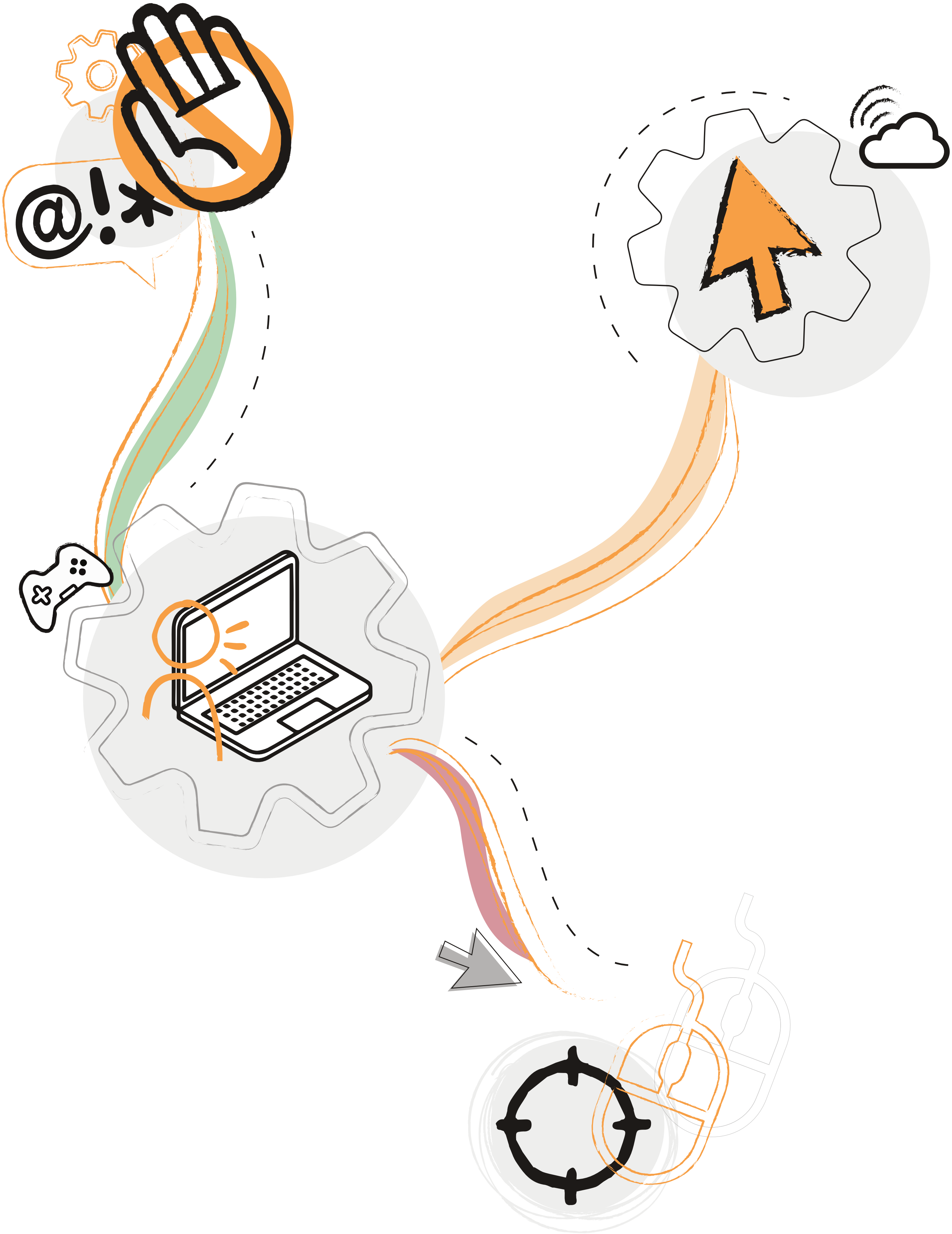


Report Harmful Content Annual Report 2022

Prepared by Prof Andy Phippen and Kathryn Tremlett for SWGfL

SWGfL Safe, Secure, Online

REPORT HARMFUL CONTENT

UK Safer Internet Centre

Contents

Table of Figures

Executive summary

This light touch analysis has considered the case file of the Report Harmful Content service from April 2021 to November 2022. During this period the service has dealt with 2,195 inquiries and has escalated around 33% of these to other services. However, in 87% of cases Report Harmful Content were able to encourage industry to successfully take action. The service also signposts a lot of inquirers to use the services offered by platforms to tackle platform specific content.

Nevertheless, major platforms are in the minority of cases, and complaints are made against a wide range of online services, which highlights the need for an independent single point of contact service, rather than assuming that signposting to major platforms is all that is required.

The analysis shows that concerns range from something that someone believes, subjectively, is unacceptable to far more serious and complex cases around illegal content and harm directed to an individual. It also shows the breadth of concern for others that can be exhibited in the case analysis.

The most common form of complaint relates to harassment or bullying, either to the inquirer or concern regarding the abuse of someone else. 754 cases were classified as harassment or bullying. Pornography is the second most common form of complaint, with 532 cases. These will generally relate to concerns around observing pornography on platforms and non-consensual posting. Impersonation – creating accounts using someone else's identity is the third most common, with 307 cases.

The data also highlights the lack of understanding about online harms in a lot of cases and the importance of providing the public with services that allow them to develop their knowledge and become more proactive in tackling concerns.

Most importantly, it highlights the importance of an independent service to signpost and direct individuals to specific resources and solutions. The service makes use of an extensive network of partners and, due to close links with industry, can both triage concerns and also more directly link concerned individuals with platforms should this be necessary.

The Report Harmful Content Service routinely refers to over 15 types of other service, the most popular (631 cases) refers inquirers to platform specific reporting pages, but they will also frequently link to other helplines, such as the Professionals Online Safety Helpline (POSH) and the Revenge Porn Helpline (RPH), police, and many other local authority services.

> — *Violent Content*
>
> *"I am in a group chat with some people from my university. One of them has sent a link to a social media profile saying, 'this is insane.' Without knowing what the link would lead to I opened it to find a video of someone being shot. I feel physically sick, and I don't know how to report it. Surely groups like this on social media aren't allowed."*

Executive summary

Analysis:

2,195
Total reports

Where this form of escalation has taken place Report Harmful Content were able to encourage industry to successfully take action in **87%** of cases

705
required additional support from other services

8% of those were escalated with industry for further action

24%
of reports escalated with industry platforms were actioned within 24 hours

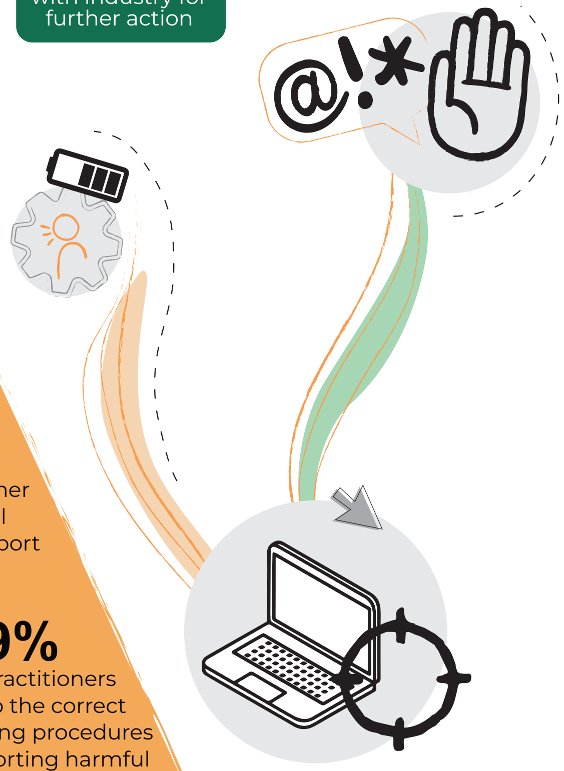
Practitioners dealt with **58%** more reports than in the last 18 months

In **52%** of cases, a reporter is making contact on someone else's behalf

33%
of reporters were signposted to other services (advocacy, civil and legal services, etc) for more bespoke support

Bullying, harassment and pornography were the most commonly reported harms, accounting for **59%** of all reports

In **29%** of reports, practitioners signposted to the correct industry reporting procedures to facilitate reporting harmful content online



Analysis of 20 months of reports received between April 2021 and November 2022

Main online harms reported:

754 Bullying & Harassment	**251** Abuse
307 Impersonation	**52** Unwanted Sexual Advances
532 Pornography	**51** Self-harm/Suicide
101 Direct Threats	**147** Violent Content

The Report Harmful Content website has been accessed via reporting buttons downloaded across the UK approximately **11,000** times

> *— Threats*
>
> *"I was raised in South Asia and am of Muslim culture. I was entered into an arranged marriage aged 18, however someone has uploaded an image of me on social media from when I was 17 holding hands with another man that is not my husband. I am extremely worried as my local community have now seen this image and have started asking my family questions as to whether I was a virgin before I got married. I have told my family that this wasn't the case, however they do not believe me and think I have brought shame and dishonour to them. People are now making threats, including my husband, saying they want to kill me, and I have been forced to leave my family home. I really need help getting these images removed.".*

About the Report Harmful Content Service

Report Harmful Content (RHC) is a national impartial dispute resolution service that has been designed to assist everyone with reporting harmful content online. RHC is provided by the UK Safer Internet Centre and operated by SWGfL. The service grew out of SWGfL's experience running the Professionals Online Safety Helpline and the Revenge Porn Helpline. Whilst these services offer essential support to members of the children's workforce and adults experiencing intimate image abuse, respectively, certain elements of online safety provision were identified, with which neither of these helplines could assist:

- Bullying & Harassment
- Pornography
- Impersonation
- Abuse
- Violent Content
- Direct Threats
- Unwanted Sexual Advances
- Self-harm/ Suicide

RHC was designed to fill that gap. It empowers anyone who has come across harmful, but not necessarily criminal, content online to report it by providing up-to-date information on community standards and direct links to the correct reporting facilities across multiple platforms. The service also provides further support to clients based in the UK, over the age of 13, who have already submitted a report to industry and would like outcomes reviewed. RHC is able to act in this mediatory dispute resolution role with a number of industry platforms, with whom it has a trusted flagger partnership and their reporting flows integrated into the RHC website. These platforms include: Facebook, Instagram, Snapchat, Twitter, Roblox, TikTok, Discord, Twitch, Pinterest, Yubo, Match Group (which includes Match, OK Cupid, Tinder, PoF and Twoo), Microsoft (which includes LinkedIn, Bing, Xbox, Skype and Minecraft) and Google (which includes YouTube, YouTube Kids, Google Search and Blogger). All dispute resolution offered by RHC is provided to clients via email contact.

About the Report Harmful Content Service

The term ‘harmful content’ can be very subjective. In order to remove ambiguity, specialist online safety practitioners studied the community guidelines of several different industry platforms. They found that eight areas of content are likely to violate platform terms: abuse, bullying and harassment, threats, impersonation, unwanted sexual advances, violent content, self-harm/suicide content, and pornographic content. Report Harmful Content (RHC) practitioners offer impartial dispute resolution associated with these eight types of online harm. They also offer advice on further issues faced online and signpost to support services and the police when necessary.

As mentioned above, RHC works in trusted flagger partnerships with a number of industry platforms. It also works closely with government departments, both in terms of designing the service and providing consultation on new policies. Due to the complex nature of online harms and their impacts, the service also maintains relationships with, and makes referrals to, other support agencies, charities, the police and social services. This report has been designed with all of these parties in mind, in the interests of information sharing for best practice. More broadly, this report will also be of interest to academics, researchers, journalists and others with an occupational interest in online safety.

— *Self-Harm or Suicide Content*

*“I used to be a self-harmer myself for many years and would often have suicidal thoughts. Luckily, I was able to get the help and support I needed. However, whilst on social media this week I have come across some concerning videos online of people actively self-harming themselves. In addition to this, the comment sections are filled with people talking about how they self-harm and how they think about suicide. I am fortunate not to be triggered by this anymore, but I am sure there are others out there who would be. Can we request this post for removal?”**

*Throughout this analysis there are examples of initial reports made to RHC. These have been lifted from client reports over the last 20 months and anonymised to maintain client confidentiality. Aside from the quotes in the executive summary and, overall, the anonymisation, the text in these quotes has not been edited and is a true reflection of the original contact made with the service. We felt it was important that the voice of those seeking support be portrayed as authentically as possible and that, by taking this approach, the complexity and impact of the issues reported would be better understood.

About this Analysis

This analysis draws upon reports handled by the service between April 1st 2021 and November 30th 2022. During this time the service dealt with 2195 reports. Each time someone makes a report to the service a case is opened, with details recorded on the nature of the concern, where the content is hosted, whether there are wider concerns, and the outcomes of the inquiry. A narrative around the case is also recorded by Report Harmful Content. The following analyses this case log in both quantitative and qualitative ways to determine the nature of the inquiries, the range of concerns and the different types of outcomes that arise from the inquiries.

The nature of reports can range from something as simple as a fact check about the legitimacy of the sort of content someone has seen, to serious cases with criminal activity that require the intervention of the police. A number of cases will be escalated to platforms with which Report Harmful Content has relationships, covering all the major social media companies, and there will also be a large number of outcomes that will require more simple intervention such as signposting to platform services or other forms of support. The service works effectively as a first point of contact for those with concerns around something that has either happened or they have seen online. In a lot of cases (1,153) an inquirer is making contact on someone else's behalf, which can also make direct intervention more of a challenge and in a lot of these cases signposting to other services is effective.

— *Impersonation*

"I have noticed that a profile has been set up in my name on social media, using my profile pictures and a URL link in the bio to a fake adult's site insinuating that I have images on there. The profile also has really degrading comments on the images of myself. I think it has been set up by my ex-partner but he has denied this. I am just worried that this account is making people believe that this is who I am and what I do. Is there anyway I can get this removed?"

As mentioned above, in total between the analysis dates the service received 2,195 reports. These are generated via email or the service's web interface. Once the initial report has been made, a case is created, and decision made regarding whether the case is sufficiently serious that it requires escalation to a platform or other partner (for example, police), or whether the inquirer simply requires support or guidance to resolve themselves. In some cases there are also general inquiries rather than a specific take-down request.

In total 705 cases had some form of escalation, 152 (8.3%) of which were to partner industry platforms to take action (i.e. remove content, apply sensitivity filters or regain access to hacked accounts). Where this form of escalation has taken place Report Harmful Content were able to encourage industry to successfully take action in 87% of cases.

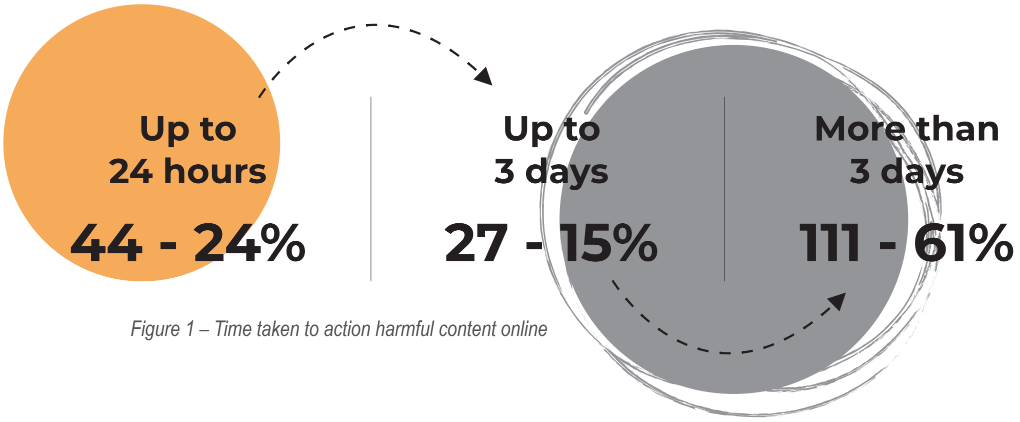


Figure 1 – Time taken to action harmful content online

This is not to say that the other reports should be dismissed as inappropriate or having no value to the service. It is important to note that as well as providing a broker service to platforms and other providers, the service also provides a support/counselling role for those who might be upset by something that is occurring online.

> *— Online Abuse*
>
> *"I regularly go to watch my local football team play, however, there have recently been a group of people who have started to shout racist insults to the black people on the team. Unfortunately, this has now become apparent on a local online group where derogatory comments are being made about some of the black players. I commented on the post to call these people out for the racist language; however I am now receiving messages from people insulting me and making derogatory remarks. I have reported it all to the platform but no action has been taken."*

While the resolution to this might not be a take down or a criminal matter, that is not to devalue the support that the service can provide towards that individual. The database is a useful resource for understanding the knowledge held by those contacting the service around online harms and, more broadly, digital literacies. As such, the data is invaluable as a snapshot of attitudes toward online harms and this is also explored in this analysis.

In total, decisions recorded following an inquiry are:

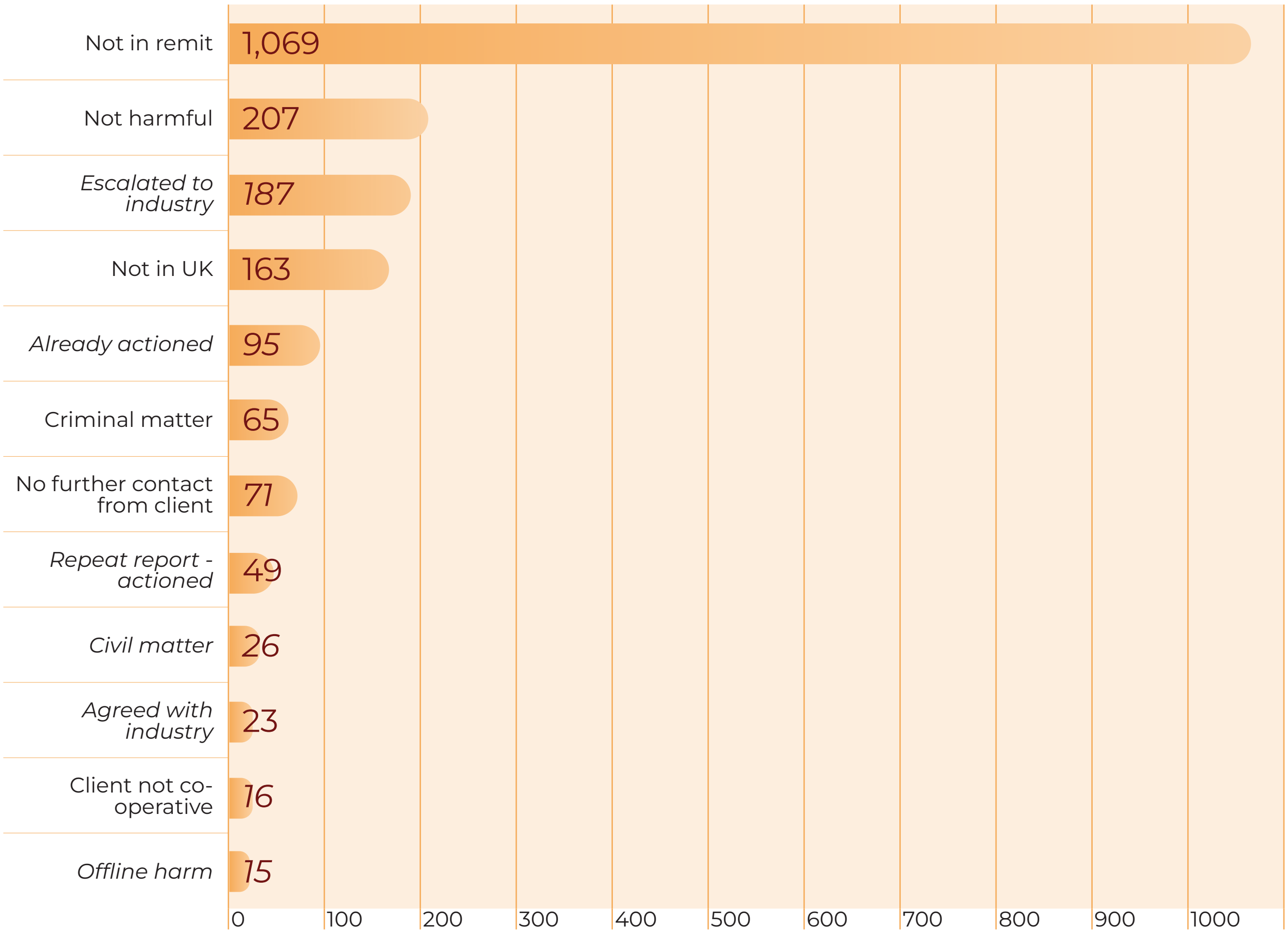


Figure 2 – Decisions recorded following reports made

> — *Pornographic Content*
>
> *“My son is 14 years old and has been looking through his explore page on Instagram. He told me that on this page he came across a woman in her 30s showing her genital area. When he clicked on her profile (as a lot of young boys would do), he found several other images of naked woman. My son has since felt nervous about using Instagram again and is worried he is going to get in trouble. I have reported it to Instagram, but they have not responded to any of my reports. Please can you help have this profile removed?”*

Those terms italicised in figure 2 refer to cases that have some form of escalation. However, in general figure 1 shows the volume of cases that are not specifically in scope, whether as a result of the fact the complaint does not relate to content that could be considered harmful, because the client did not engage further or they were not in the UK.

Nevertheless, in a lot of these cases, the service would still provide inquirers with details on how to take further action themselves, signposting specifically to platform reporting.

If we consider the breadth of referral to other services:

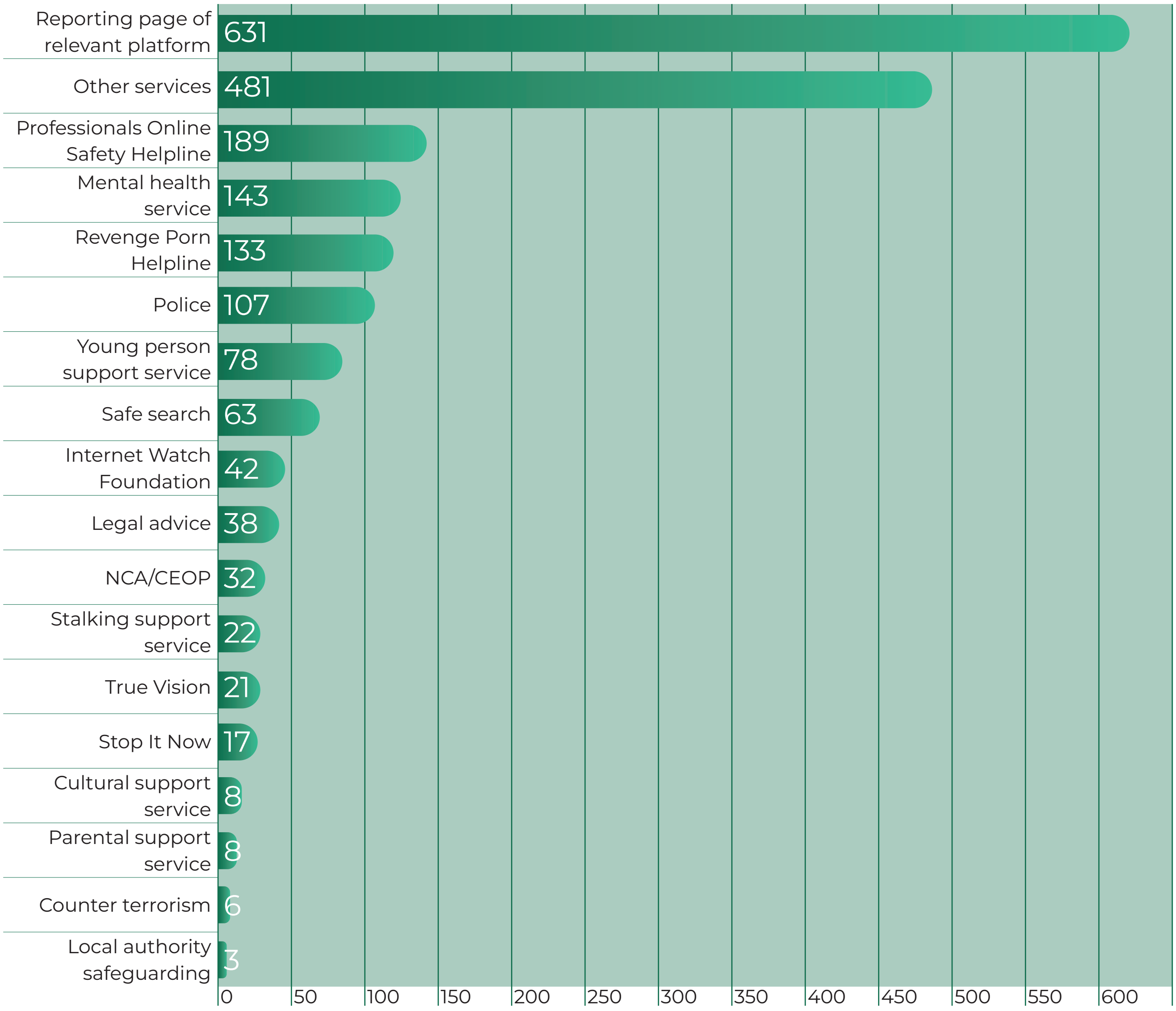


Figure 3 – Referrals to other services

We can see that the service works within an extensive network of other support services. In the event of some types of content, such as Child Sexual Abuse Material (CSAM), the service will defer to the specialist services of the Internet Watch Foundation. The service also works closely with other helplines provided by SWGfL, such as the Revenge Porn Helpline and the Professionals Online Safety Helpline. Furthermore, in a large number of cases, they will direct the inquirer to platform specific reporting. "Other service" can be a wider range of options, such as local authority support and other services in an inquirer's local area.

Another type of data recorded by the service relates to the platform where the "harm" is taking place. This is, again, useful data because it shows occurrence of harms and, in subsequent analysis, any changes in platforms used for harms. While the platform frequencies are not surprising, there are a large number of "other":

390	TikTok	**258**	Facebook	**254**	Instagram
73	Twitter	**70**	YouTube	**39**	Snapchat
30	Google	**18**	Discord	**7**	Roblox
6	LinkedIn	**3**	Xbox Live	**2**	Blogger
2	Twitch	**1**	Match.com	**1**	Tinder
1	Bing	**1**	Minecraft	**637**	Other

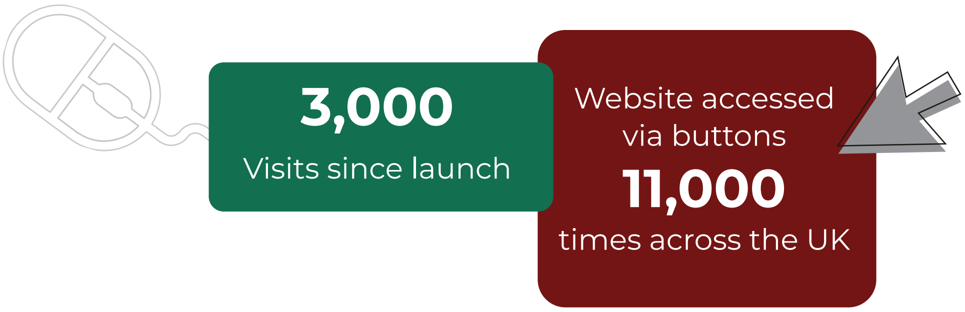
Figure 4 – Breakdown of harmful content by platform

"Other" can relate to direct contact, for example via messaging platforms or mobile communications, but also many other minor platforms and services. While there are few surprises that Facebook, Instagram and TikTok are all key platforms for harms, there is also a great deal of diversity in this data. If we are to aggregate this data, the "major platforms" are in the minority of calls. Which, again, highlights the need for an independent single point of contact service, rather than assuming that signposting to major platforms is all that is required.

Reporting Button

In October 2021, Report Harmful Content introduced a new innovative and accessible way for users to report legal but harmful material online through the Report Harmful Content button, a quick and simple method for helping anyone to report offensive or harmful material online, no matter where they are. The button has been developed to offer anyone living in the UK a simple and convenient mechanism for gaining access to reporting routes for commonly used social networking sites, gaming platforms, apps and streaming services alongside trusted online safety advice, help and support. It also provides access to an online mechanism for reporting online harm to the RHC service for those over the age of 13 where an initial report has been made to industry but no action has been taken. RHC will review content in line with a site's community standards and act in a mediatory capacity where content goes against these.

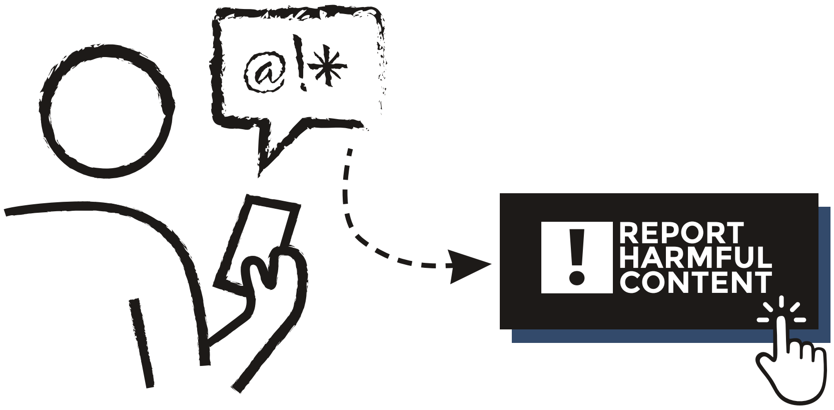
Since its launch, the downloads page on the website, which contains the code and instructions for embedding onto websites, has been visited nearly 3,000 times and the reporting pages on the Report Harmful Content website have been accessed via buttons across the UK over 11,000 times. The fast adoption of the reporting button by many secondary schools across the UK and the resulting support this offers pupils, staff and the wider community has been of particular interest. In order to encourage and empower more people to report harm online, we would recommend that this initiative be promoted widely across the UK through media literacy strategies across the DCMS and DfE.



3,000
Visits since launch

Website accessed via buttons
11,000
times across the UK

Figure 5 – Reporting button usage figures



Reiya

Whilst analysing data about the times and days reports were made to the helpline in 2022, practitioners observed that a large amount of reports were being submitted outside of operating hours, in the evenings and at the weekends (accounting for 33% of the reporting case load between Jan 2021 and Feb 2022). Similarly, a not insignificant portion of reports were coming from people residing in countries other than the UK. One of the services sister helplines, The Revenge Porn Helpline also observed a similar pattern for cases they were responding to. To help provide a solution both helplines worked together to create a Chatbot, Reiya which could provide support at the time of searching through advice and signposting with the option to triage this as a case and pick up in operating hours if needed.

Launched in February 2022, Reiya had 9003 sessions and 64,989 interactions up until the 30th November 2022. Of these, pornographic content, online abuse and bullying and harassment were amongst the top 6 issues selected. The proportion of reports made out of hours to the service over the weekend has dropped since the launch going from 33% between January 2021 and February 2022 to 20% during February and November 2022 showing that Reiya is having the desired effect in reducing out of hours reporting, providing advice and signposting in the moment of need.

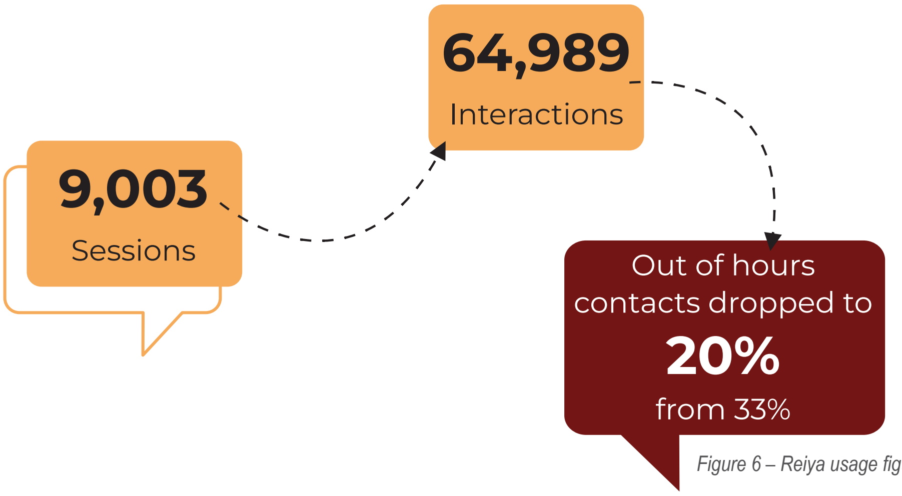


Figure 6 – Reiya usage figures

Content/Behaviours

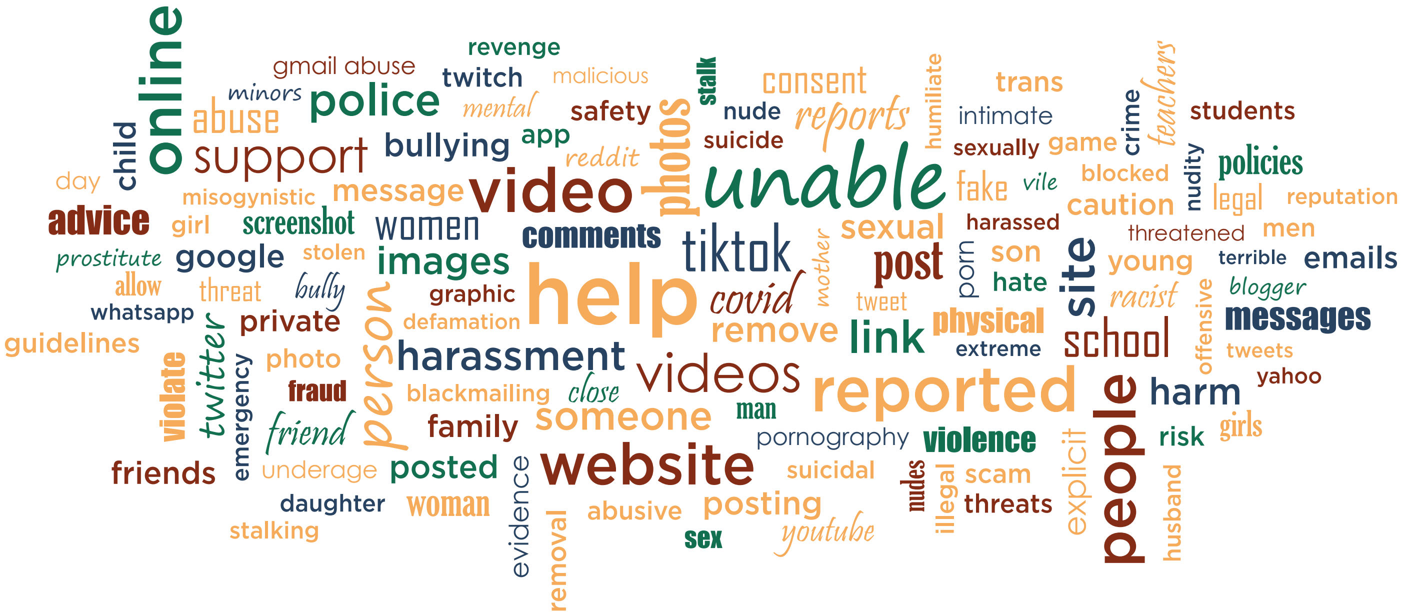


Figure 7 – Wordcloud highlighting the nature of reports made to the service

The above word cloud highlights the nature of reports to the service. A word cloud is a simple way of looking at the frequency of words in a body of text – in this case an aggregation of all of the initial reports to the service. In this snapshot we can see that the more widely used words are "help" and "unable" – highlighting that this is a service that is used by people that need help with upsetting content, but also that they have, perhaps, been unable to get help from other services and platforms. We also see "reported" and "support" as significant words, again highlighting the need to help. We can also see different types of behaviour (such as bullying, harassment and consent), as well as different platforms and content types being strongly shown. However, it is perhaps most interesting to see "people" and "person" also coming to the fore, reminding us that these harms are caused in general by others – while they might manifest on online platforms, the root cause is generally another person.

Each report will be categorised, which means we can analyse the prevalence of different types of harms. While, as discussed above, this does not mean that every call will meet a sufficient threshold for further escalation, it does allow us to explore the nature of the concerns of reporters to the service:

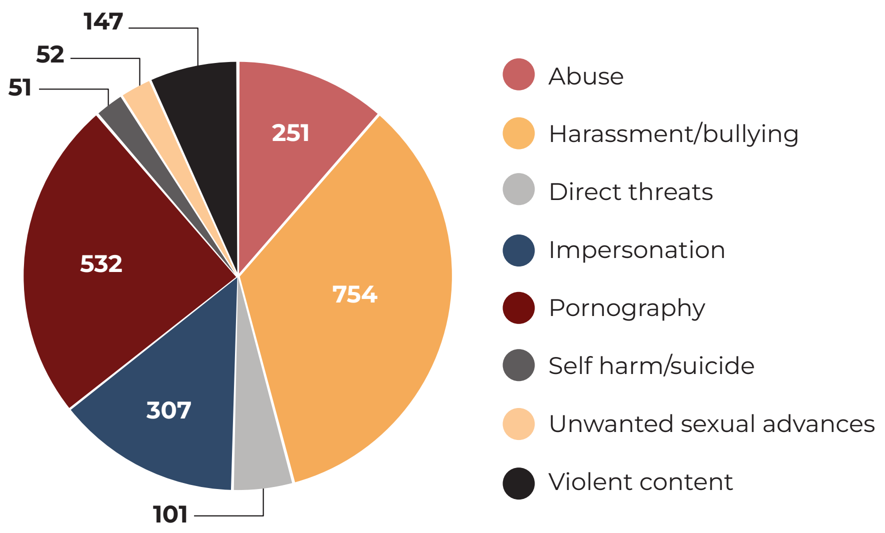


Figure 8 – Breakdown of reports by primary type of harm

Content/Behaviours

The following draws illustrative comments from cases to show the nature of each of the categories, ordered in terms of prevalence. These are illustrative quotes, and it should be noted that each categorisation will have many other different types of complaint. The quotes are used to highlight the breadth of concern and the nature of reports the service receives.

Harassment and Bullying

We can see from this data that Harassment and Bullying is the most prevalent complaint made to the Report Harmful Content service, which categorises prolonged abuse of an individual. For example:

> *I want my image taken down from this page, I have reported it over an over again it is harassment and Instagram have not removed it for image privacy. I have reported to police that ask this is removed. I also have various crime reference numbers. Please can you escalate this as I have reported numerous times and I want my image removed from this harassment post, it is bullying. My full name is being revealed with my address.*
>
> *I am emailing about an account which is relentlessly harassing transgender people and trans allies. He recently created a post dedicated to harassing a trans woman, he consistently refers to them as male, published a picture of them as a child (pre-transition, therefore humiliating her) without her consent, and makes numerous libellous and {{username}} accusations. The Equality Act 2010 protects transgender people (and those perceived as or associated with transgender people) from harassment which describes actions which humiliate, offend or degrade people based on their being transgender.*

Pornography

This relates to any content that someone might consider to be pornographic and will generally relate to complaints either about a website someone has seen, or the actions of an individual on a wider platform:

> *Good afternoon, I wanted to inform you that your platform is hosting illegal content on several forums. These are images depicting celebrities in a state of nudity or engaged in acts of sexual conduct created and posted without their permission, including depictions that have been faked. Also be aware that teenagers access and comment that content.*
>
> *This person is impersonating me on social media and created a porn website to scam people. I haven't been able to report the website because I've been blocked from it and from the Instagram profile. However my followers can see everything that's going on*

Content/Behaviours

Impersonation

Specifically relates to people creating fake accounts and the nature of complaints generally relate to those being impersonated wishing to have accounts removed or taken down:

> *My ex husband has been harassing me - this is currently with the police as a live case. He has created fake profiles of me across all social media platforms. The profile picture is a photo of me that I sent to him when we were married. I reported the profile to LinkedIn, they took it down and then allowed it to remain!*
>
> *I have already reported the account that is using non-consensual intimate images of me using my name and sending requests to friends trying to scam men for money but Instagram hasn't taken the profile down. I would highly appreciate your help on that matter. Thank you in advance*

Abuse

The Abuse category is quite broad, and can relate to a wider ranging of content that inquirers are referring to as either harmful or upsetting. It can be a challenging category to address because the interpretation of harmful can be extremely subjective:

> *Anti-vax book. Amazon seems to be a hotbed for Rightwing conspiracy theory books. Absolutely dreadful and dangerous.*
>
> *I reported the account for depiction/promotion of weed and vaping but there was no response from platform.*
>
> *It is in Italian, but I understood it, and the last comment is written in English. A man posts pictures of another man's feet, claiming he makes photos of them and touches the feet, while the other man is sleeping. I find it disturbing.*

Violent content

Violent content is also a broad category that can be subjectively interpreted by those making reports and can be concerned about the implications for others seeing the content:

> *I stumbled onto this video, and have not watched it, but the still image looks very disturbing and it is called 'dog crush'. I very much hope it isn't filled with violence towards animals, but it very much sounds like it is.*
>
> *The website is full of posts making direct and indirect threats towards women and young girls. I didn't dare open half of the things on there, I became aware of it having seen someone share a post saying their pictures were used on it. The website talks about raping women and uses people's images without their consent, the website is vile and should be removed.*

Direct threats

Direct threats are a narrower category than some of the others, and will relate to specific threats made to an individual or group of individuals:

> *I am being threatened by people who published my pictures of me and my family and they threaten to kill us. Please delete these posts from their personal pages on Facebook. I send the addresses of the links in which the posts are located. Please review them and delete these posts and my pictures of me and my family. To the Jews, they considered us traitors to the country. We fled the country, but we are under threat. Please delete all photos and posts. I cannot communicate with Facebook.*
>
> *In January this year, I received the blackmailing sextortion email. that I ignored and reported to Google. after few days I received another e-mail with my name added to the sexual context mailing with porno pictures, which I ignored and reported to google. After that I started to have a massive attack in Google search with my name added to phishing and malware software running in the browser, as well with direct threats with back-linking my name taken from another platform I think that someone specially added my address in the data base, I repeat again for the fraud action with direct threat, sexual harassment. I reported the case already to the police*

Self harm/suicide

Another specific category is self harm/suicide – in a lot of these cases it will, again, be individuals concerned with content they have seen and how it might impact upon others who might see it:

> *Fake information about covid19 treatments. Driving young people to do dangerous untested vaccine shots. Comments blocked on the website.*
>
> *I {{name}} be wrong about what I want to report, and I'm really scared that the person who I want to report be dragged into something when it is completely unnecessary. I'm just asking, what happen regarding my report? What do you do when you receive a report about a potentially suicidal person?*

Unwanted sexual advances

And finally, the unwanted sexual advances (not image based) category can relate to both advances to the individual and also concern for others:

> *Gaydar hosts many chatrooms some of which contain messages from people who are clearly interested in under-age sex. The Phone Sex room is the main one but there are others. I have complained to Gaydar about this in the past but they haven't replied.*
>
> *Pegi 16 on Google but Pegi 12 on apple. This is on my 12 year old daughters phone. The AI is very sexually and dominatrix submission style. It is not acceptable and I don't believe it should be acceptable for even 16 year olds. The higher the level you get the more graphic it gets. It's apparently an AI best friend for teens with depression!!!*
>
> *I have suffered months of harassment on Tiktok. I have naked pictures and pictures in the hands of someone I don't know. That's why it threatens me. He says he send the photos to my family if I don't do what he says. Although I reported the person to Tiktok. His account is still active. And he's contacting me.*

Content/Behaviours

As stated above, these samples should not be considered exhaustive – they have been selected to show the breadth of inquiry and what people consider to be online harms and harmful content. It is useful to highlight the importance of having a central, informed, service that can support those who have concerns, given the broad nature of concerns, and also the vast network of referral that the service can offer. By way of final illustration, the nature of the complaints are also categorised against "wider issues", that record other factors beyond the core harmful content concern. These are:

Abuse/Malpractice	134	Cultural/Religious Sensitivity	29
Malicious Communication	18	Coercive Control	4
Mental Health Concerns	44	Hate Speech	83
Child Sexual Abuse Material	47	Extortion	40
Grooming/Child Sexual Abuse	27	Hacked Account	61
Rape/Sexual Abuse	29	Harassment	70
Intimate Image Abuse	107	Credible Threats	18
Voyeurism	4	Animal Abuse	13
Fraud	11	Searching for Harmful Content	109
Identity Theft	10	Stalking	20
Domestic Abuse	15		

Figure 9 – Wider issues reported in addition to primary harm

Specific mental health issues can also be recorded:

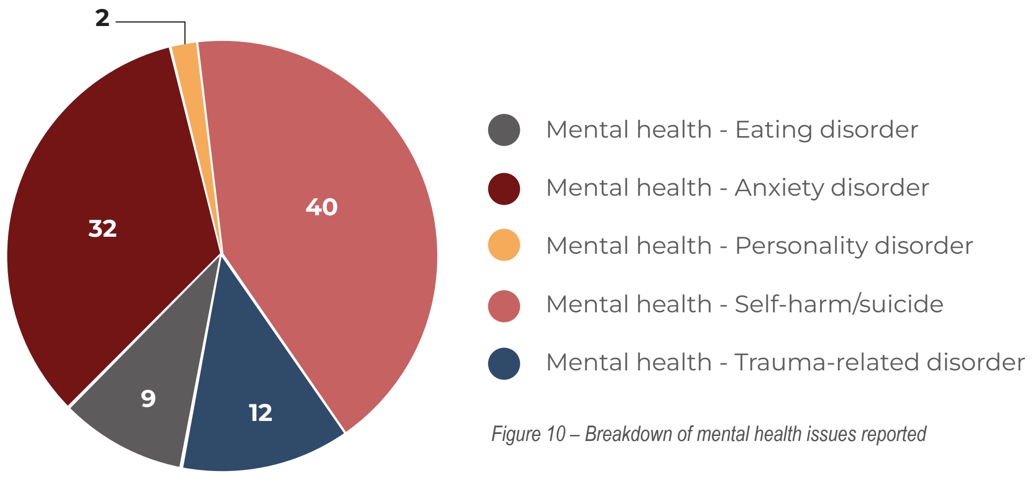


Figure 10 – Breakdown of mental health issues reported

A lot of these wider issues relate to offline concerns and, in a lot of cases, illegal activity or behaviours that require some form of professional intervention which, again, are unlikely to be supported without a centralised, knowledge drive, service being in place. Almost half of the reports recorded in the last year have some form of "wider issue" that might require further specialist intervention.

Implications

This light touch analysis has considered the case file of the Report Harmful Content service from April 2021 to November 2022. During this period Report Harmful Content has dealt with 2,195 reports and has escalated around 33% of these to other services – this might be specifically reaching out to an industry partner to take action on their platform (removal of content, applying sensitivity filters or regaining access to hacked accounts), but equally it might be a referral to another service (such as the Internet Watch Foundation in the event of illegal child abuse content) or other forms of support. The service also signposts a lot of inquirers to use the services offered by platforms to tackle platform specific content.

The analysis shows that concerns range from something that someone believes, subjectively, is unacceptable to far more serious and complex cases around illegal content and harm directed to an individual. It also shows the breadth of concern for others that can be exhibited in the case analysis.

This also highlights the lack of understanding about online harms in a lot of cases and the importance of providing the public with services that allow them to develop their knowledge and become more proactive in tackling concerns.

Most importantly, it highlights the importance of an independent service to signpost and direct individuals to specific resources and solutions. The service makes use of an extensive network of partners and, due to close links with industry, can both triage concerns and also more directly link concerned individuals with platforms should this be necessary. Without such a centralised service, there would be considerable duplication of effort across platforms and, in the case of the many smaller platforms and peer to peer abuse, potentially leave those concerned or vulnerable with no place for support.

> — *Unwanted Sexual Advances (Not Image Based)*
>
> *"I had an online relationship with someone a few months ago, however I broke this off when I didn't feel comfortable with some of the conversations we were having and how intimate he wanted to be online. I am now getting sent unwanted nude images from someone on social media, every time I block this contact, they come back on another account and send another image. Their face isn't in the images, but I have a feeling it could be the same person. Is there anyway I can make this stop?"*

Implications

During the last 12 months the UK Government has seen unprecedented turbulence resulting in continued delays to the Online Safety Bill (OSB) progressing through parliament and receiving royal ascent. At the time of writing this report, the OSB stands to dismantle an essential obligation that supports victims of online harm. Current Video Sharing Platform Regulation requires platforms in scope (e.g. Twitch, Snapchat and TikTok) to provide an impartial procedure for the resolution of disputes between users and industry providers. As it stands the OSB is set to remove these existing crucial obligations that impartial or independent arbitration of complaints remains to support users and victims of online harm.

The joint select committee reported in December 2021 encouraging the Department for Culture, Media and Sport to look towards Report Harmful Content as a potential model for what an ombudsman could look like concluded that it is only through the introduction of an external redress mechanism that service providers can truly be held to account for decisions as these are what impact people.

Report Harmful Content allows people this redress route and has reported the impact of legal but harmful content on victims online since launching in 2019. The report you have just read is further evidence emphasising the importance of an impartial appeals process. Crucially, had the service not been here, harm occurring online may not have been realised or addressed.

Resources Discussed

Report Harmful Content

South West Grid for Learning

The Professionals Online Safety Helpline

Report Harmful Content button

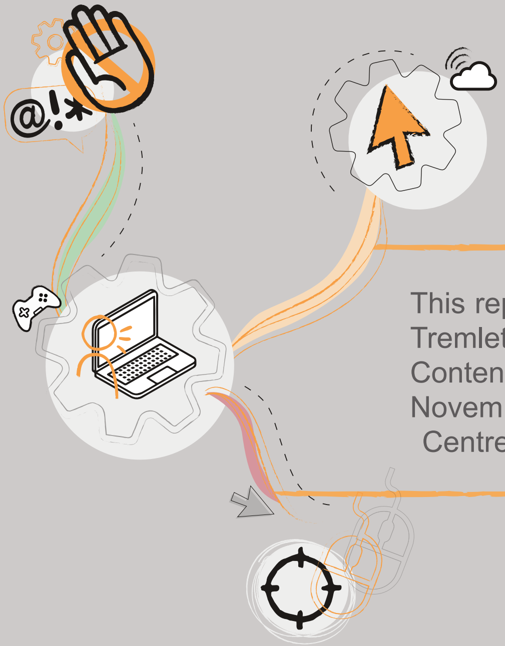
The Revenge Porn Helpline

Internet Watch Foundation

Action Counter Terrorism

Reiya

Video Sharing Platform Regulation



This report has been prepared by Prof Andy Phippen and Kathryn Tremlett reflecting on the reports made through the Report Harmful Content service during the 20 months between April 2021 and November 2022. Report Harmful Content is a UK Safer Internet Centre service provided by SWGfL and 50% funded by Nominet.

SWGfL
Safe, Secure, Online

REPORT
HARMFUL
CONTENT

UK Safer
Internet
Centre

NOMINET