Ofcom

making communications work
for everyone

# Your response

Please refer to the sub-questions or prompts in the annex to our call for evidence.

| Question | Your response |
|---|---|
| **Question 1: Please provide a description introducing your organisation, service or interest in Online Safety.** | 1. Cifas is the UK's fraud prevention community. We are a not-for-profit membership organisation which leads the fight against fraud by sharing data, intelligence and learning. We have over 30 years of experience in fraud prevention and financial crime, working with a range of UK businesses, charities, and public bodies to help protect themselves, their customers and the public.  At the time of writing there are over 630 organisations in Cifas membership, and a full list of Cifas members can be found on our public website. We are a Specified Anti-Fraud Organisation (SAFO) designated under the Serious Crime Act 2007 as recognised by the UK Government.<br>2. We provide dynamic intelligence and cutting edge fraud prevention tools to understand the fraud threat landscape and strengthen our community against fraud. We deliver accredited learning and trusted training for organisations and individuals. Cifas data feeds into the National Fraud Intelligence Bureau (NFIB) to aid the disruption and investigation of fraud and wider organised crime and is officially counted within the ONS crime statistics.<br>3. The Cifas National Fraud Database (NFD) is used to share information about fraudulent conduct against member organisations. We call this fraud |

| | |
|---|---|
| | risk information. An example of this is when someone uses another person's details, such as name and address, to apply for a credit card. The NFD prevents over £1 billion in fraud losses every year and protects the public from abuse of their identity to obtain financial and public services.<br>4. The fraudulent conduct shared by our members is so often facilitated through the abuse of online channels, including large social network and search engine platforms. This ranges from social engineering on social media to share personal and financial information, to the trade in data and documents through sites available at the click of a search engine button. These details and documents are then used to obtain or hijack services in the name of innocent members of the public.<br>5. Abuse through these channels also includes money mule recruitment through posts and videos on major online platforms which, at its most insidious, is helping to enable the sale of guns and drugs, people trafficking and human slavery.<br>6. The enablement of fraud and wider organised crime and the significant harm this causes to both victims and wider society makes effective regulation of major online platforms so vital. The regulatory requirements of the Online Safety Bill (OSB) represent a significant opportunity to reduce the facilitation of fraud and wider crime through these major online platforms, if effectively implemented. |
| **Question 2: Can you provide any evidence relating to the presence or quantity of illegal content on user-to-user and search services?** | 1. In the points below we have summarised key threats in relation to abuse of those online channels to enable fraud and wider crime. This summary has been drawn from analysis of data held by Cifas, intelligence shared between our |

**IMPORTANT: Under this question, we are not seeking links to or copies/screenshots of content that is illegal to hold, such as child sexual abuse. Deliberately viewing such images may be a criminal offence and will be reported to the police.**

members and law enforcement partners and of open source material.

2. Abuse of social media and online platforms can be broadly categorised as follows
   i. Trading and exchange of personal information to support subsequent fraud.
   ii. Sharing of fraud tactics, expertise and methodologies.
   iii. Recruitment style activity to identify and entice individuals to engage in fraud
   iv. Scams style activity and promotion of fraudulent or fraudulently purchased products. This includes employment scams. For example, where jobs don't exist, and victims are duped into making payments for fake employment checks or courses.

3. In addition to the financial losses and harm caused directly to victims, frauds enabled through these platforms both fund and facilitate the most serious crime types, including terrorism – as evidenced within this independent RUSI Research**.**

4. There are some key contributing factors which combine to make the scale and resulting harm of this abuse so significant
   i. 96% of UK households have internet access - ONS
   ii. More than three quarters of the UK population are now active on social media, with 80% of those users actively engaging in the past month - Talkwalker
   iii. Which? research identified that one in ten people have been scammed by adverts on social media or search engines - 45% claimed they would not be able to tell if an advert was fake or not

iv. The lack of effective verification of social media accounts enables criminals to hide behind the anonymity of fake profiles when targeting victims, and in some instances impersonate trusted sources

v. The range of messaging and communications options available through these sites enables approaches to victims at scale, in an instant

5. There are some key themes around the nature and harm from this threat

   i. Fraud and other criminal activity are intertwined, and these activities are often perpetrated by the same individuals or organised criminal gangs.

   ii. The same social media platforms that facilitate cyberbullying, abuse and the promotion of exploitative content are also being used to facilitate fraud and promote fraudulent goods and services.

   iii. The sale of illegal goods and services through the sites are often an outlet for money laundering to facilitate organised crime.

   iv. Money mules - those who allow their accounts to be used to launder the proceeds of crime, including fraud - are often recruited online through adverts (including paid for ads) on social media for "get rich quick schemes". Social media is used by organised gangs to groom and corrupt students and vulnerable adults, in order to entice them into this criminality.

   v. Romance Fraud is often facilitated through social media and forms a rich source of income for fraudsters, many of whom evade capture by committing the crime from abroad, using websites accessible within the UK.

| | |
|---|---|
| | 6. Effective, water-tight regulation of user-to-user and search services has the potential to disrupt complex fraud networks, including those engaged as part of wider organised crime. |
| **Question 3: How do you currently assess the risk of harm to individuals in the UK from illegal content presented by your service?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 4: What are your governance, accountability and decision-making structures for user and platform safety?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 5: What can providers of online services do to enhance the clarity and accessibility of terms of service and public policy statements?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 6: How do your terms of service or public policy statements treat illegal content? How are these terms of service maintained and how much resource is dedicated to this?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 7: What can providers of online services do to enhance the transparency, accessibility, ease of use and users' awareness of their reporting and complaints mechanisms?** | *Is this response confidential? – Y / N (delete as appropriate)* |

| | |
|---|---|
| **Question 8: If your service has *reporting or flagging* mechanisms in place for illegal content, or users who post illegal content, how are these processes designed and maintained?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 9: If your service has a *complaints* mechanism in place, how are these processes designed and maintained?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 10: What action does your service take in response to *reports* or *complaints*?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 11: Could improvements be made to content moderation to deliver greater protection for users, without unduly restricting user activity? If so, what?** | **This response provides our answer to questions 11 and 18 because they are strongly inter-linked.**<br>1. There is an opportunity to learn from the 630 Cifas member organisations which span financial services, wider private, public and third sectors, who all conduct appropriate fraud and identity verification checks as part of the consumer journey. This includes at the point of application for a product or service.<br>2. Online platforms must be required to move from reactive and retrospective action based on consumer reporting, to proactive and effective due diligence checks. For example, the verification of advertisers before their adverts are published, which can be achieved with minimal disruption to genuine advertisers or the public and prevent the significant harm from fake or misleading adverts. For example, adverts for jobs that don't exist and are, in fact, a way to dupe victims into transferring funds |

through their accounts and unwittingly commit money laundering.

3. Cifas members conduct Know Your Customer (KYC) and identity verification checks on each and every application, utilising Cifas and other external and internal data points. This is integrated within the customer journey, including those products where a very quick decision on approval is required. In short, the vast majority of business passes these checks and is then assessed against the usual criteria, with referrals made for the minority of applications where further investigation is required. This ensures minimal disruption to genuine parties while identifying fraudulent conduct, across millions of applications. The use of KYC and identity verification checks could and should be mirrored in the checks required of these online platforms.

4. The fraud identified through those checks will create an opportunity for analysis and learning that can be blended into those checks, including red flags on recurring features of fraudulent behaviour. Indeed, there is scope for a range of technology based tools to be deployed to add further resilience, as part of a holistic and honed anti-fraud response, as seen already in financial services and other sectors. For example, the use of machine learning and behavioural analytics. In addition to checks on profiles and advertisers, this approach must also be used to proactively identity and take down the fraudulent content posted by those criminals operating under a range of fraudulent accounts, under different names.

5. Automation can absolutely play a role in this, but it can’t be on the basis of best endeavours or provide the whole solution, or fraud will be missed, and innocent parties disadvantaged. For example, if a previously impersonated advertiser or user is simply blocked, when in

| | |
|---|---|
| | that instance they are a genuine party putting forward legal content. Automation must form part of an effective end to end process to identify fraudulent applications, accounts and content, with appropriate points of human intervention, including for investigation and consumer redress.<br>6. There should also be a requirement for information and data sharing; not just with law enforcement, but between the platforms. When one Cifas member identifies and shares fraudulent conduct through Cifas, all 630+ Cifas member organisations are instantly protected and also able to protect current customers and those genuine future applicants. This type of sharing should be mirrored by online platforms to ensure that all platforms and their users have the same protection from fraudulent content once identified by a single platform. |
| **Question 12: What automated moderation systems do you have in place around illegal content?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 13: How do you use human moderators to identify and assess illegal content?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 14: How are sanctions or restrictions around access (including to both the service and to particular content) applied by providers of online services?** | 1. As outlined in our response to earlier questions, there is too often a reactive approach from the platforms to fraudulent conduct and those behind it. Too often the same content can reappear under a different account, run by the same person(s), both on that platform and a range of others. It is not sufficient to simply take down a post or revoke a single account. |

| | |
|---|---|
| | 2. This again highlights the critical importance of platforms conducting checks through a range of technologies and data points, and the non-competitive sharing of data and intelligence across platforms. This is crucial to proactively identifying fraudulent content and the full range of accounts behind it, and taking effective action against those accounts. Only with this proactive and structured approach to checks and data sharing can platforms deliver meaningful disruption to criminals abusing their platforms, and better protect consumers and wider society from those threats and harms. |
| **Question 15: In what instances is illegal content removed from your service?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 16: Do you use other tools to reduce the visibility and impact of illegal content?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 17: What other sanctions or disincentives do you employ against users who post illegal content?** | *Is this response confidential? – Y / N (delete as appropriate)* |

| | |
|---|---|
| **Question 18: Are there any functionalities or design features which evidence suggests can effectively prevent harm, and could or should be deployed more widely by industry?** | **Please see our response to question 11, which provides our answer to questions 11 and 18 because they are strongly inter-linked.** |
| **Question 19: To what extent does your service encompass functionalities or features designed to mitigate the risk or impact of harm from illegal content?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 20: How do you support the safety and wellbeing of your users as regards illegal content?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 21: How do you mitigate any risks posed by the design of algorithms that support the function of your service (e.g. search engines, or social and content recommender systems), with reference to illegal content specifically?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 22: What age assurance and age verification technologies are available to platforms, and what is the impact and cost of using them?** | *Is this response confidential? – Y / N (delete as appropriate)* |
| **Question 23: Can you identify factors which might indicate that a service is likely to attract child users?** | *Is this response confidential? – Y / N (delete as appropriate)* |

| | |
|---|---|
| **Question 24: Does your service use any age assurance or age verification tools or related technologies to verify or estimate the age of users?** | *Is this response confidential?  – Y / N (delete as appropriate)* |
| **Question 25: If it is not possible for children to access your service, or a part of it, how do you ensure this?** | *Is this response confidential?  – Y / N (delete as appropriate)* |
| **Question 26: What information do you have about the age of your users?** | *Is this response confidential?  – Y / N (delete as appropriate)* |
| **Question 27: For purposes of transparency, what type of information is useful/not useful? Why?** | *Is this response confidential?  – Y / N (delete as appropriate)* |
| **Question 28: Other than those in this document, are you aware of other measures available for mitigating risk and harm from illegal content?** | *Is this response confidential?  – Y / N (delete as appropriate)* |

Please complete this form in full and return to OS-CFE@ofcom.org.uk