Ofcom

Consultation on Additional Safety Measures

Response from the Commissioner Designate for Victims of Crime for Northern Ireland

3rd October 2025

1. General Comments

1.1 The Commissioner Designate for Victims of Crime for Northern Ireland welcomes the publication of Ofcom's proposed additional measures to further strengthen online safety for UK internet users. The Commissioner notes that this work is part of Ofcom's year of action on online safety and that the proposals sit within a broader suite of law, codes and policies to protect and safeguard people online. Positive action to make the online world safer for our citizens, in particularly our children and young people, is to be welcomed. It is encouraging that Ofcom is endeavouring to take all action possible to safeguard UK internet users within the confines of the Online Safety Act and European Convention on Human Rights.

1.2 The Commissioner Designate recognises that many of the proposals outlined, in this consultation and in the Codes and Online Safety Act, rely on technological innovations of which this office does not have the requisite expertise to evaluate. As such, this response does not seek to assess the technical feasibility or efficacy of the proposals, rather it will focus on the needs of victims and the value of the proposals in preventing harm.

1.3 The most effective strategy to improve online safety necessarily includes both preventative measures, which stop people from being victimised in the first place, and effective responses when it is found that someone has been targeted. The Commissioner Designate is assured that the proposals outlined focus on both prevention and intervention. More concrete obligations for online service providers to build early intervention into the fabric and design of products will be key to making safety a reality.

1.4 Whilst the Commissioner understands the need for measures to be proportionate and feasible, there is a risk that the caveats relating to technical feasibility within the measures outlined may enable service providers to sidestep their responsibilities. One example that has already arisen in the context of the Online Safety Act is service providers stating they are unable to access or police messages where end-to-end encryption is a component of the service. This is an instance where service providers may be considered limited in their commitment to safety by design. It is not lost on the Commissioner that the financial sector has developed systems that are very secure yet have functionality for institutions to access systems to investigate fraud. If such a balance can be struck in financial systems, there

is surely scope to develop systems which enable services to effectively monitor illegal content. Development of such systems to make services safer for users and safeguard against illegal content being published should be the cost of doing business in the UK. Service providers should have to show willingness to seek technical solutions that provide similar user benefits but which meet the safeguarding needs of our most vulnerable. This is especially the case for larger services with broader reach, significant revenue, and the ability to pivot to new ways of working if required.

1.5 The Commissioner Designate welcomes the commitment at paragraph 1.54 to continually review the effectiveness of measures. The pace of technological innovation, and its tendency to constantly evolve and outpace the mechanisms put in place to monitor legal compliance, remains a huge challenge. An iterative approach allows for agility to change and improve processes to ensure that measures are as effective as they can be.

2. Livestreaming

2.1 The Commissioner Designate welcomes the additional measures outlined to make livestreaming safer for children and young people.

2.2 The requirement for a mechanism to report harmful or illegal livestreamed activity, and for human moderators to review content and take action in real time, are key. Moderators may be more likely to assess more subtle forms of abuse or coercion and be alive to risks relating to terrorism, incitement to hatred or suicide. It is important that moderators have the capacity to review content swiftly and stop livestreams if deemed to be harmful or illegal.

2.3 We acknowledge that the live nature of livestreaming can pose particular challenges for moderators. The proposals outlined provide a logical and effective means of meeting those challenges. The introduction of an ‘imminent physical harm’ reporting category with expedited processes governing the response by service providers in event of such reports, would be a positive measure. It is vital that in the event of a live crisis, such as a gun attack or incitement to violence against a group of people, service providers also have routes to provide law enforcement with information as required.

2.4 It is encouraging that Ofcom has devised further solutions to protect children who livestream, and that it has been recognised that more needs to be done given the difficulties associated with moderation of live content in real time. The proposals to prevent reactions, comments, gifting or content capture on livestreaming by children are proportionate and necessary steps to safeguard children online.

2.5 The addition of highly effective age assurance is a welcome step to bring alignment within the Codes by updating the Children's Code to incorporate these measures. The use of highly effective age assurance is a positive step to tackle grooming and mitigate the risk of exploitation of young people online.

2.6 It is concerning that there remain gaps within these provisions, including accounts run and owned by an adult which live streams content involving a child, and the omission of overseas accounts accessed by UK users from the proposals. Whether this is a matter for Ofcom or for legislators, it is unacceptable that Ofcom is not deemed to have a role in cases where adults access an online service within the UK to watch live streamed child sexual abuse if the abused child is outside the jurisdiction. Any service which hosts such content on a UK-based service should be held accountable and should be covered by obligations to prevent such content from being hosted on their site and take action in the event that it is.

3. Proactive Technology

3.1 The Commissioner Designate supports the proposals to utilise proactive technology to identify child sexual abuse material.

3.2 It is not for this office to assess the efficacy of proactive technologies. However, as such technology develops it will be important for a human element to remain to monitor such technology's efficacy in real life cases, in particular cases where there is nuance or uncertainty. The Commissioner Designate therefore welcomes that human review is a component of these proposals.

3.3 Whilst we recognise Ofcom's desire for measures to be agile so as to apply broadly, there is a risk that the measures may not be prescriptive enough to create concrete obligations to use proactive technologies to safeguard users.

3.4 We would also query whether proactive technologies could be utilised where feasible to pre-screen content at device level, thus preventing harmful or illegal content from being posted or shared in the first place. This approach would be a more effective safeguard than removal after harmful or illegal content has already been published.

3.5 It would also be useful if such proactive detection technologies could be employed to detect acts of hate and stalking given that these inherently contain a risk of escalation and threat to safety. If assessed that there is a risk of violence, service providers should utilise reporting channels to inform the relevant law enforcement agency under crisis management protocols.

4. Hash Matching

4.1 The Commissioner Designate is encouraged that additional measures to increase the use of hash matching technology is included in this proposal. As we stated in our previous response to Ofcom on a safer online for women and girls:

Given the prevalence of attacks on women and girls online, including child sexual abuse, exploitation, domestic abuse and stalking, practices such as red teaming and hash matching should also be a core component of how online service providers operate to stress test their services for abusability. It should in particular be mandatory to stress test in relation to the ability of users to commit criminal acts, or to use platforms to facilitate future criminal behaviour, such as grooming children and moving them to private or encrypted platforms before committing a crime.

4.2 It is welcome that Ofcom proposes to use these technologies for CSAM, terrorism and intimate image abuse content. Consideration should be given as to whether such technologies might also be usefully employed in cases where content is shared to coordinate attacks on a certain section of the community but not under a specific banner of a proscribed organisation. This would be especially helpful in cases of racial unrest, such as those following

the Southport attacks and the reported sexual assault of a young woman in Ballymena Northern Ireland, which triggered days of riots and targeting of ethnic minority residents in the area, including sharing locations of people attempting to shelter from the violence. The key flashpoints of unrest and role of social media and online services in coordinating this violence and spreading of disinformation has been examined in CAJ's *Mapping Far Right Activity Online in Northern Ireland* report[1]. Although the recommendations in the report were focused on the responsibilities of devolved government as opposed to Westminster and its wider regulatory role, it demonstrated clear evidence of the harm that can be done to individuals and wider society if proactive and innovative efforts are not made to tackle use of online spaces to stir up hatred and violence. Consideration should therefore be given as to whether this kind of proactive technology might usefully be employed to prevent the spread of video and photo-based messaging that aids the coordination of such attacks or identifies individuals for the purposes of encouraging targeted violence.

4.3 It is also worth noting that in its final paragraph, the CAJ report issued the stark reminder that *"the harms of the identified racist discourse on social media are unfolding in a context where racist attacks and paramilitary-linked intimidation are already widespread in Northern Ireland".* This unique context should be borne in mind when considering any measures that are applicable to Northern Ireland and in assessments of risk by service providers. In particular, the report indicates a grey area whereby much activity may not be officially sanctioned by a proscribed organisation, yet there are likely elements with links to those organisations who are helping drive this online activity and broader unrest. It would therefore be helpful if the definition of terrorism content as it applies to this section and to Ofcom's regulatory work is flexible enough to incorporate such nuance.

5. Recommender Systems

5.1 The additional measures outlined to prevent recommender systems from amplifying illegal content is welcome. Removing flagged content from recommender systems until it has been moderated is a sensible approach

[1] PRINT Rabble report

which mitigates the potential harm of such content until it can be properly reviewed.

5.2 Given that Ofcom is being non-prescriptive regarding how this might work, it is vital that the efficacy of this measure is monitored so that measures can be tweaked and made more prescriptive if needed.

6. User Sanctions

6.1 User banning is likely to be an effective measure and have a significant impact on user behaviour if implemented correctly. In particular, the risk of a permanent user ban introduces an important level of friction that is likely to modify or temper user behaviour. The Commissioner agrees that a complete user ban is appropriate for users who share CSAM. For motivated perpetrators, a time limited ban may only result in a user waiting out the ban then resuming this harmful practice.

6.2 Given that people sharing CSAM often operate across multiple platforms, it would be useful for services to be obligated to communicate with one another and flag banned users to encourage banning across all platforms.

6.3 Ofcom's proposal of a varied approach for adult and child users who share CSAM is helpful, particularly in cases of children who receive content because they are being abused and children who share materials without full cognisance of their actions. As young people are more likely to be capable of desistence given the right interventions, and not always fully aware of the implications of their harmful behaviours, it may be disproportionate to implement a permanent ban for those under 18. A two or three-strikes policy may be more appropriate to enable children to learn and change, and it is appropriate to enable discretion in child cases. It may be useful to develop a framework to assist service providers in making such assessments, in conjunction with experts in the CSA field, to ensure effective action is taken in cases of motivated child perpetrators such as Alexander McCartney where levels of harm and malicious intent are significant.

7. Crisis Response

7.1 The Commissioner Designate strongly supports proposals to require service providers to draw up a crisis response protocol and to introduce dedicated communications channels for use by law enforcement in the event of a crisis. These measures will enable swift action to be taken in the event of a crisis such as the aftermath of the Southport attacks and repeated outbreaks of organised anti-migrant violence across Northern Ireland. There is evidence that online tools, including public social media and private online messaging, are being used to foment racial hatred and coordinate racist violence in Northern Ireland – this is having a demonstrable impact on the safety of our newcomer communities and on our societal cohesion. Such impacts highlight the influence of online communications in escalating and de-escalating such tensions, and the importance of online service providers doing their part in maintaining peace and order in our society.

7.2 It is important that service providers develop protocols that are capable of responding to such issues however they present. Research[2] by the Committee for the Administration of Justice (CAJ) in Northern Ireland, outlines how user patterns to organise and escalate violence may have some unique elements in Northern Ireland, including linkages to paramilitarism and longstanding sectarian divides. Findings also highlighted how local anti-migrant campaigning and unrest didn't necessarily link in with broader UK-based far right networks and messaging in a way that happens elsewhere.

7.3 It is also important that, once a crisis is de-escalated, law enforcement is able to request and receive evidence from online service providers to facilitate prosecutions. This is part of a wider issue, in which increasingly evidence of crime is on U2U platforms and is required in a format that can be entered into evidence to prosecute a case in criminal court. Local law enforcement and legal professionals continue to tell this office that they experience significant delays in receiving materials requested from service providers to facilitate criminal prosecution. Of even greater concern is that local police have highlighted cases where they are unable to get responses from social media companies to take down harmful or dangerous material, including doxing materials. Social media companies must be more responsive in such cases, particularly as failure to do so can escalate into a crisis situation where life is at risk. Dedicated communication channels with

[2] See PRINT Rabble report

adequately staffed departments with knowledge and authority to work with law enforcement should therefore be able to be utilised to both prevent a crisis and to allow justice processes to proceed without delay in the aftermath of a crisis.

If you would like to discuss any of these points in further detail, please contact the office via:

Tel: 028 9052 6607
Email: policy@cvocni.org