Online Harms White Paper

Response from the Mayor of London

Summary

- The Mayor of London welcomes the opportunity to comment on the government's Online Harms White Paper. New technologies can absolutely be a force for good, but the regulation and legislation that currently exists has not kept pace with the development and growth of the internet, and more must be done to ensure that our online world is a safe and positive place for UK citizens to explore.

- The Mayor's Police and Crime Plan for London has committed to taking a zero-tolerance approach to hate crime and he is acutely aware of the growing impact of the internet in the spread of hate speech, extremist views, and the harassment of democratic representatives. The Mayor and the Deputy Mayor for Policing and Crime have been vocal about the role of the internet in spreading violent messages and the incitement to commit serious youth violence.

- The Mayor's Office for Policing and Crime (MOPAC) has worked very closely with a range of partner organisations and with social media providers. There has been positive action within the industry to work collaboratively, to improve their understanding of these issues and to develop new approaches to proactively manage harmful content on their platforms. However, the progress made by some of the major technology companies does not disguise the harm already caused. Good will and self-regulation are not robust enough on their own to protect the public from the wide range of harms apparent across the internet. Therefore, the Mayor welcomes the Government's proposal to establish a new statutory duty of care to make companies take responsibility for the safety of their users and tackle the harm caused by content or activity on their services.

- The Mayor believes that a framework of comprehensive regulations overseen by an adaptable new regulator is the way to provide a standardised set of rules and expectations for the public, technology companies, and government alike. This must include a statutory code of practice to tackle content which may although may not currently be illegal, would be recognized as harmful to individuals, harmful to our democracy, or against fundamental British values.

- No business or industry should ever consider itself above the national rules or laws set by democratic processes. Social media platforms already have a legal obligation to remove content that breaks local laws. But this is not always happening or happening quickly enough. With the skills and resources these companies have at their disposal, we agree it is possible for them to go further and faster. The onus must be on platforms to clean their own houses, and for law enforcement to only be involved in obvious and serious criminality.

- The Mayor appreciates however, that there must be a balance between content intended to bully, intimidate, or humiliate, and that which seeks in good faith to openly challenge and criticize ideas and institutions. Regardless of what legal powers may exist, there will always be content that does not meet the criminal threshold but might still be incredibly harmful, and there has to be a framework in place for tech companies to proactively deal with these circumstances.

- Tied into new codes of practice and robust regulation must be a culture of transparency and accountability. The public who use these platforms should be aware of the prevalence of harm, and it is right that any regulator is able to draw upon a wide range of interventions against companies who allow harmful content to spread on their platforms, or who refuse to adopt the UK's regulatory requirements.

- It is encouraging that some of the major tech companies have welcomed these proposals, and that they recognise the clarity and the benefits that clear rules and guidance can provide to them, but it is imperative these proposals are not diluted to become ineffectual. Comprehensive action is needed, and it is already overdue.

Transparency

- The Mayor agrees that accompanying new codes of practice and regulation must be a culture of transparency and accountability. This will help to show where action is being taken to improve safety online and ultimately improve trust in some platforms. We are supportive of the measures set out in the white paper and the proposal to utilise enforcement action where necessary.

- It is especially important that the public have confidence that a platform's use of algorithms, AI and machine learning are not exerting bias towards online harms. The regulator will of course need to balance the reporting of the use of these mechanics with protecting proprietary information, but it is clear there are worrying examples of platforms pushing users to ever more extreme and harmful content[1]. This is recognised in the White Paper, but the government must ensure the regulator has enough power to act against companies whose platforms continue to covertly steer users towards this content. General corporate ethical codes on the use of algorithms, machine learning and AI may not on their own be sufficient to reassure the public that sufficient safeguards are in place. The regulator must be assured that companies have robust procedures which provide actionable guidance internally to avoid harm.

Scope of Online Harms

- The Mayor is pleased to see the wide range of online harms recognised within the White Paper. This reflects many of the concerns we have around the safety of users online, from terrorism and child protection to bullying, harassment and disinformation. It is also reassuring to see the broad definition for companies in scope of the regulatory framework, which would appear to cover the clear majority of current and future platforms.

- The omission of private communications from the scope of these proposals is understandable, however, current social media giants are looking to move away from the present open, 'public square' experience towards emphasising the use of private channels

[1] Home Affairs Select Committee – Hate Crime and its Violent Consequences – Oral evidence session 24 April 2019: Transcript at http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/home-affairs-committee/hate-crime-and-its-violent-consequences/oral/100660.pdf

for communication and content sharing.[2] Although this move may help to restore trust in platforms rife with harmful content, it poses a problem for the future efficacy of the proposals within the White Paper. There must be a clear, fair, and consistent definition developed over what truly constitutes private communications in the online space to prevent these regulations becoming meaningless.

- We do not have a set view on what number of participants in a communications channel would qualify as public or private, however, there should be no safe space online for the promotion, planning, or facilitation of terrorism, serious violence or child sexual exploitation or abuse (CSEA). Even within private communications, there should be a function for users to report harmful behaviour to the platform to act within the agreed code of conduct.

- Any definition of private communications should not wholly hinge on whether initial entry to a platform or group is by invite. For example, there is a clear difference between a one-to-one chat via a communication-specific platform - such as WhatsApp - and a one-to-one message on a platform whose main purpose is for entertainment, such as an online video game. This second example can allow harmful actors, such as paedophiles, a way to engage vulnerable children against a backdrop of an innocent gaming activity.[3]

Victims and User Redress

- The Mayor recognises that in recent years cooperation with major tech companies has improved. However, moderation and reporting processes for online harms, such as hate crime and harassment, are inconsistent, opaque, and with little to no feedback for the user. Trusted Flagger status schemes have helped to improve how some harmful content, such as inflammatory gang videos and hateful content, is identified and removed, but it cannot be left to civil society organisations and our already stretched law enforcement agencies to police all of this content. We would like to see the regulator set out guidelines for how tech companies will more proactively monitor and remove this content, and how they should empower users to report online harms and signpost them to support. Ultimately, much of this content should not be permitted to be uploaded in the first place.

- The Mayor's Violence Against Women and Girls (VAWG) strategy recognises the impact that online abuse and harassment can have. A recent study conducted by Opinium for the children's charity Plan International UK showed nearly double the number of girls (23%) said they felt harassed regularly by someone through social media, compared with 13% of boys. Harassment ranged from unwanted contact, trolling, and cyberbullying to sexual harassment and threats of rape and murder. These figures are backed up by the results of the 2018 Youth Survey, which also indicates a worrying proportion of children have been exposed to other harms such as discriminatory or gang-related content.[4]

- There is also serious gap in our response to types of sexual offences which are image based. Sexting and so called 'revenge porn' including the use of 'deep fakes'[5], fall within this

[2] https://www.nytimes.com/2019/04/30/technology/facebook-private-communication-groups.html
[3] https://www.mirror.co.uk/news/uk-news/paedophiles-using-online-computer-games-10233554
[4] Pages 22-24 https://www.london.gov.uk/sites/default/files/youth_voice_survey_report_2018_final.pdf

[5] https://metro.co.uk/2019/05/31/deepfake-porn-ethics-able-watch-whatever-imagination-desires-9526079/

category. These are offences that are overwhelmingly committed against women and the current legal position means that it is a challenge for authorities to address perpetrators in line with broader forms of sexual offending. The Mayor welcomes the government's recent announcement to review the law around the non-consensual taking, making and sharing of sexual images.[6]

- Online abuse and hate crimes cover the spectrum of bigotry towards minority groups, from Islamophobia and Anti-Semitism, to transphobic and homophobic prejudice, and mistreatment of Roma, Gypsy and Traveller communities. Feedback from partners in 2015 showed that whilst only 2-5% of hate crime reported to the Metropolitan Police Service concerned online cases, the Community Security Trust evidenced that around 20% reports of anti-Semitic hate crime were online, and TellMAMA reported up to 75% of the Islamophobic cases they received were online. This demonstrated the huge gap between what was reported to the police, and what was being reported via third parties. Online hate crime is known to have multiple victims as it can affect any community or individual who identifies with the victim's group or community, and it spreads extremely quickly beyond city and country borders.

- MOPAC has committed to ensuring its commissioned support services are equipped to meet the needs of victims of online offences, but there is far more that can be done. The proposed duty of care must put the onus on the platform owners to enforce the codes of practice against users who create or share harmful content, and not on the victims of this content to merely block or mute the offenders. The regulator will need to educate users, especially the young and vulnerable, as to what their rights are around online harms, and to provide oversight of the complaints processes companies have to build user confidence and encourage reporting.

- As previously stated, many users of these platforms have started to see online abuse as an everyday occurrence and therefore do not report these incidents. The regulator should play a key role in developing ways for platforms to improve the support they provide so that women and girls have the confidence to report this abuse. A key element of this is evidence gathering and it is essential that technology companies work with partners to improve this. Technology companies need to be willing to provide the technological solutions to retain and disclose relevant and required evidence in a way that doesn't disadvantage the victim and is appropriately captured to meet the required threshold for the CPS to support criminal cases. There has been reluctance from the technology industry in the past to support this aim and this needs to change for more prosecutions to happen.

Offenders and Enforcement

- The Mayor welcomes the proposal to establish requirements for referring clearly illegal content to law enforcement and other relevant government agencies to aid investigations, and in the case of the most harmful content, such as terrorism and CSEA, this should be done proactively by the technology company and that processes for preserving evidence of offences are streamlined as much as possible. For harms which are not clearly in the criminal

[6] https://www.gov.uk/government/news/law-around-non-consensual-taking-making-and-sharing-of-sexual-images-to-be-reviewed

space, it is important the regulator can direct technology companies to take proactive action to deal with those who breach the codes of practice. The police do not have the resources to deal with a mission creep which may result from large numbers of referrals which are not clearly of a criminal nature.

- To assist with ensuring these platforms remain safe, we support the regulator to ensure companies do more to hold individual users to account for their actions online. This would include processes to share intelligence of those most harmful, repeat offenders with other platforms, so that there is a unified and consistent response across the internet. It is not merely enough to remove harmful content and channels, the users themselves but be held to account by companies in a robust and consistent manner. This is especially true of users who fundraise across platforms by espousing hateful and 'non-violent' extremist views and users which try to hide behind online anonymity in order to perpetrate harm against others.

- The White Paper proposes a range of enforcement options for companies which fail to fulfil their duty of care, and we welcome these, especially the suggested powers to impose significant fines, senior management liability, and ISP blocking for the most serious and continued of breaches. However, tech companies, the public, and law enforcement will need more clarity as to what constitutes legal and illegal harms when committed online. Most legal provisions in this area predate the era of mass internet and social media use. For example, legislation such as s44 and 45 of the Serious Crime Act 2007 are no longer enough to tackle the growing problem of online videos being used to inflame and promote gang violence. A review is needed to modernise our criminal law to better deal with serious online harms.

Design

- The Mayor welcomes the proposal for government to develop a Safety by design framework to help companies incorporate online safety throughout the development or update of their online services. As we have seen with the crime prevention initiatives of Secured by Design, (an organisation overseen by MOPAC) planning and designing products at an early stage with safety and security in mind creates an overall securer environment and prevents expensive changes at a later date. This could lead to a Safe by Design accreditation which companies could use to promote their sites and provide confidence to users. This is especially important for platforms which either direct their services towards or are used regularly by under 18s.

- A regulator should provide information, advice and support to tech companies that are less successful than the major companies in blocking and removing the most harmful content, therefore building capacity across the tech sector to disrupt the promotion of terrorist and CSEA material.

Education

- It is clear that as a society we face a bigger undertaking than just the important task of providing citizens with better digital skills. We need more digital understanding, which focuses on the ability both to use technology and to comprehend, in real terms, the impact that it has on our lives. We urge the government to support initiatives and partnerships with business and civil society on online safety, privacy, and critical thinking - to empower users to stay safe online for internet users of all ages.

- The government and the regulator should also work to ensure this education is proactively pushed out to adults as well as children to guarantee they are not left behind by the rapid advancements in technology. But more broadly, we should be ensuring that we support children and young people to learn more about the diverse communities they live in and are part of so that we are tackling many of the causes of hate crime. People are less likely to engage or be caught up in divisive rhetoric if they have a greater sense of belonging in their local communities and engage more positively with their neighbours and others in their local area. Prioritising work on improving social integration at a national level, in line with the work the Mayor is promoting in London, would be a good way to tackle this

> Education is a key strand to the prevention of online harms. In December 2018 the Mayor and Google jointly announced the provision of a £600k grant from Google.org for charities to train youth workers to be confident in dealing with issues relating to social media and enable young people to use social media for good. But what is really needed is a step change in the way our society is taught about the internet and the

Conclusion

- The benefits to our lives from the internet are huge, but in a short space of time our online world, aided by the lure of anonymity and the absence of physical proximity to others, has become a conduit for some of the worst facets of human behaviour. We would accept this in no other area of our society, and it is time our democratic and tolerant values were enshrined in law to apply to online content.

- Wide-ranging and impactful regulation is needed and the proposals within this White Paper are a welcome start, but the work to protect against online harms must accelerate before further harm is done to the vulnerable, to our shared values, and to the foundations of our democracy itself.

For more information, please contact Leigh Greenhalgh, Principal Government Relations Office on 020 7983 4147 or at leigh.greenhalgh@london.gov.uk