Opening Statement by Karlin Lillington, Irish Times journalist to the Grand International Committee on Disinformation and Oireachtas Joint Committee on Communications, Climate Action and Environment

7 November 2019

My name is Karlin Lillington. I've been a technology journalist and columnist for over 20 years, primarily with the Irish Times, and for several years in the past, the Guardian. My particular interests are the ways in which technology interacts with and impacts society, politics, culture, the law, and human and civil rights.

I am grateful to the Committees for this invitation to offer a perspective on the troubling issue of online harms, hate speech and electoral interference. Having covered the technology sector since the 1990s, and used the internet since the 1980s, I have watched and experienced the gradual transformation of this powerful communications medium from a small, non-visual virtual space that required specific written screen commands to access and use, to the vast and varied global multimedia platform of today.

The internet I wrote about in the 1990s is nothing like the internet of today, for both better and worse. Like many who were on the net early on — and even despite being a fan of dystopian fiction — I imagined an almost entirely positive future in which the new World Wide Web would fulfil the galvanising and uniting global commons envisioned by so many early adopters of Sir Tim Berners-Lee's creation. But, so many of us failed to fully comprehend the problems that could and would also arise — risks to security and privacy, the bullying and threats, the eating away at the foundations of democracy, the elevation of the mob, rather than the commons.

All of these problems have been exacerbated, explosively, by the arrival of social media and search platforms that combine worldwide reach with personal data gathering at unimaginable, interlinked and mostly invisible scale, obscured behind non-transparent, constantly evolving algorithms. These are written by a mostly youthful, white male workforce of alarmingly low diversity, drawn from computer science departments that often have no required privacy or security course content[1], much less mandatory teaching on ethics, or on removing bias in coding. In addition, these companies all operate on the business model of 'surveillance capitalism', the surreptitious intake and exploitation of our data, and increasingly, our free will, documented in detail by Shoshona Zuboff's monumental recent study, The Age of Surveillance Capitalism[2].

Even 'data', as a term, erases the fact that it comprises the very essence of us — our likes and dislikes, our physical and emotional attributes, our social connections, our physical environment, the patterns of our daily lives. Data is us, packaged and sold on for monetisation and further exploitation by third parties and data brokers most of us have never heard of.

The best way to elicit ever more of our collectable 'data exhaust' — once so-called because it was seen as the useless byproduct of our online activity before the platforms discovered it could be lucratively repurposed AS the product — is via ever more engagement with the platforms, and hence, algorithms that by design favour the clickbait material of hate, of outrage, of conspiracy, of tribalism, of fury. And from there, it's a short goosestep to threats and intimidation, electoral and political manipulation and control, the targeting and suppression of dissent. From threats, to violence, arrest, torture and even death.

---

[1] Lillington, Karlin: "Colleges fail to teach design of secure code: RSA conference hears top computer science programmes in the US deliver minimal content on security" https://www.irishtimes.com/business/technology/colleges-fail-to-teach-design-of-secure-code-1.1705828

[2] See Zuboff, Shoshona: The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, Profile Books, 2019.

Perhaps this seems a luridly extravagant or sensationalist view. But this is the very disturbing reality in many parts of the world that, I suspect, are not routinely discussed when lawmakers and experts gather to consider such problems as we are here to discuss today. This grim situation also consistently fails to be addressed sufficiently by the platforms themselves. And many of the proposed solutions and interventions — such as the suggestion that online anonymity be banned, or account registrations be tied to formal identity documents — only further these problems, rather than fix them.

Therefore, I would like to use my time here to offer evidence of some of the global impacts of online harms, hate speech and electoral interference, as they relate to the human rights defenders that are amongst those most seriously harmed, and sometimes, shockingly, murdered, as a result of activity intrinsic to all the major online platforms.

If we better understand, and more adequately address these serious harms — the way they come about, how the platforms are utilised, the worrying support sometimes given by platforms that are intent on increased user engagement and which ignore the context of that engagement, the toxic business models that cannot help but support such activity — we can better resolve the problem for all of us.

Why? Because first, and critically: online platforms are extremely important for pro-democracy activists and journalists, especially in the most repressive countries. Committee members will know how central they were to events such as the Arab Spring. Daily, the platforms are a major tool OF democracy— because they allow anonymity, because they can be used through virtual private networks to disguise location, because they have a presence nearly everywhere, even in the locations of some of the worst regimes, because some offer easy to use, encrypted messaging and phone calls.

This is why they are also exploited as tools AGAINST democracy, and why so often, the online attacks and manipulations made on human rights defenders soon become the model for attacks and exploitations closer to home. Equally, the actions taken, or not taken, by the platforms and the leaders in established democracies, can either curtail or encourage online harm. Activists are generally the first victims of experimentation by state actors, the target of organised trolls and bots, and the deliberate amplification of disinformation. They suffer the consequences of inaction or insipid responses by those with authority and control — including the platforms.

According to Front Line Defenders (https://www.frontlinedefenders.org/), a Dublin-based international NGO which works to protect human rights defenders around the clock, and which has a special expertise in digital security, evidence indicates that activists such as human rights defenders, trade unionists, and journalists draw more threats, and those threats are followed up on more often, than in the normal hate speech spiral. Andrew Anderson, executive director of Front Line Defenders, says:

> “A threat isn’t just a threat. It’s the warning signal that worse is going to happen, if it's not followed up on. There’s precedence in using these [activist] groups as indicators. This is why we should be focusing on human rights defenders.”

While hate speech and disinformation campaigns are cruel and destructive to any recipient, the risk to activists in repressive states can be life threatening, because they deliberately create an environment for far worse to happen, and be tolerated. In 2017, Front Line Defenders analysed data on the killing of 312 activists. In 84 per cent of these cases where adequate information was available, the defender had previously received a threat, many of which were made online. Ed O’Donovan, head of protection at Front Line Defenders, notes:

> “Attacking the legitimacy and credibility of human rights defenders through these platforms lessens the reaction when they are arrested. The ground has been softened.”

This approach is particularly well documented in the Philippines. False accusations and associations set the stage for the public to be desensitised to arrests or deaths, if victims were

supposedly involved in such questionable associations all along (it is estimated that 7,000-12,000 Filipinos have been killed to date by Duterte's vigilante anti-drugs gangs). Facebook is regularly used for coordinated online attacks, because Facebook, through a special programme promoted in the Philippines, is often the only way Filipinos can access the internet.

Filipino activists and journalists have identified multiple ways in which the Duterte government uses social media platforms to spread disinformation about activists and journalists, in an attempt to discredit and threaten. Maria Ressa is a former CNN correspondent who founded a news site called Rappler in the Philippines in 2012, a site which has been highly critical of the government. When Ressa received the Gwen Ifill Press Freedom Award last year, she warned in her acceptance speech:

> "This is an existential moment for global power structures, turned upside down by technology. When journalists globally are under attack. When power structures are shifting. Our problems are partly caused by yours: American social media technology platforms, once empowering, now weaponized against journalists, activists and citizens, spreading lies across borders…
>
> We need to hold tech platforms to account. They need to move away from just business growth — they are now the world's largest distributor of news so they have to take on the responsibilities journalists had as gatekeepers. They cannot allow lies to spread. They need to protect the public interest — and the public sphere where democracy happens."

Much of Ressa's commentary focused on Facebook, which she referred to as "essentially our Internet."[3] But Facebook also helped create the Duterte regime, offering training sessions to his campaign in how to use to platform, and referring to him at one point favourably in internal communications as an "undisputed king" of adept usage of the platform, according to an extensive Bloomberg investigative story in 2017"[4]. Bloomberg journalist Laura Etter wrote:

> "Repressive governments originally treated Facebook, and all social media, with suspicion—they saw how it could serve as a locus for dissidents, as it had in the Arab Spring in 2011. But authoritarian regimes are now embracing social media, shaping the platforms into a tool to wage war against a wide range of opponents—opposition parties, human-rights activists, minority populations, journalists.
>
> The phenomenon, sometimes referred to as "patriotic trolling," involves the use of targeted harassment and propaganda meant to go viral and to give the impression that there is a groundswell of organic support for the government. Much of the trolling is carried out by true believers, but there is evidence that some governments, including Duterte's, pay people to execute attacks against opponents. Trolls use all the social media platforms—including Twitter, Instagram, and YouTube, in addition to the comments sections of news sites."

Organised disinformation is a global business. In August, Facebook announced it had removed hundreds of accounts that were targeting users in Sudan with messages supporting military generals whose troops had massacred pro-democracy activists in Khartoum last summer.

---

[3] Bump, Philip: "The warning offered by a Filipina journalist targeted by her country's president" https://www.washingtonpost.com/politics/2018/11/21/warning-offered-by-filipina-journalist-targeted-by-her-countrys-president/

[4] Etter, Laura: "What Happens When the Government Uses Facebook as a Weapon?" https://www.bloomberg.com/news/features/2017-12-07/how-rodrigo-duterte-turned-facebook-into-a-weapon-with-a-little-help-from-facebook. Also see: "Duterte May Attack Our Report, But Momentum To Protect Filipino Defenders Is Building" https://www.globalwitness.org/en/blog/duterte-may-attack-our-report-momentum-protect-filipino-defenders-building/ and Mihm, Henry, Ines Oulamine, Fiona Singer: "The Philippines Deserves More From Facebook" https://www.lawfareblog.com/philippines-deserves-more-facebook

Sudanese activists "were unsurprised to learn of the campaign" associated with two Middle Eastern companies, the New York Times reported. The story noted:

> "'There have been so many fake accounts,' said Mohamed Suliman, a Boston-based engineer allied with Sudan's protest movement. 'Fake news is a real source of danger for Sudan. If there is ever a counterrevolution, one of the regime's main tools will be social media'."[5]

Similar misuse of the platform in Assam — and a criticism of Facebook's failure to adequately monitor or tackle the problem — are highlighted in a new report by human rights group Avaaz. Alaphia Zoyab, senior campaigner at Avaaz, said in a statement:

> "Facebook is being used as a megaphone for hate, pointed directly at vulnerable minorities in Assam, many of whom could be made stateless within months. Despite the clear and present danger faced by these people, Facebook is refusing to dedicate the resources required to keep them safe. Through its inaction, Facebook is complicit in the persecution of some of the world's most vulnerable people."[6]

In Pakistan, "65 per cent of all cybercrime complaints lodged in Karachi relate to harassment over Facebook, according to data available with The Express Tribune," states the annual report "Pakistan's Internet Landscape 2018" from Pakistani human rights group Bytes For All.[7]

The report also noticed campaigns of online harassment, utilising video posted to YouTube, and attacks involving posts on Facebook and Twitter.

Human rights NGO Global Witness's 2018 report "Enemies of the State" documents how many repressive governments follow an established pattern of attacks on activists that begin on social media, to legitimise harassment, arrests, imprisonment and even executions.[8]

---

[5] Walsh, Declan and Rashwan, Nada: "'We're at War': A Covert Social Media Campaign Boosts Military Rulers" https://www.nytimes.com/2019/09/06/world/middleeast/sudan-social-media.html

[6] Dixit, Pranav: "Facebook Failed The Rohingya In Myanmar. Now It May Be Repeating Its Mistakes In Assam" https://www.buzzfeednews.com/article/pranavdixit/facebook-failed-the-rohingya-in-myanmar-now-it-may-be. On Facebook and the Rohingya, see "Report of the independent international fact-finding mission on Myanmar" https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_64.pdf. Also see: Sombatpoonsiri, Janjira: "Manipulating Civic Space: Cyber Trolling in Thailand and the Philippines: https://www.giga-hamburg.de/en/system/files/publications/gf_asien_1803_en.pdf

[7] "Pakistan's Internet Landscape 2018" https://bytesforall.pk/sites/default/files/Internet%20Landscape%20Report%202018.pdf and see: "2018 saw greater controls on online free speech, threats to data privacy in Pakistan: report" https://www.dawn.com/news/1475665

[8] "In Bangladesh, legislation enacted in September 2018 imposes a ten-year jail sentence for online posts which 'ruin communal harmony or create instability', and a 14-year sentence for using digital media to 'intimidate people and/or cause damage to the state'. Dozens of people have reportedly been arrested in Nicaragua after the government introduced a new law widening the definition of terrorism in July 2018. It is being invoked against protesting students, farmworkers and other demonstrators. In Egypt, a new media regulation law allows anyone with a social media account with more than 5,000 followers to be placed under government surveillance, making them vulnerable to prosecution for posts deemed to be "fake news". Passed in June 2018, Vietnam's new Cybersecurity Law requires internet companies such as Facebook and Google to set up offices in Vietnam and store private user data which could potentially be used for state surveillance. The law has been widely criticised for the risks that it would be abused to stifle political dissent, which could include land and environmental activism." From: "Enemies of the State" https://www.globalwitness.org/documents/19766/Enemies_of_the_State.pdf, p.29.

Such actions involve all social media platforms, of course. Twitter is also plagued with hate speech, fake accounts, and coordinated bot and troll accounts. In Pakistan, Twitter was the forum for attempts to manipulate the recent national election:

> "A 'Trends Monitor' report noted that hate speech connected to the elections had taken over a large part of the online space, with a significant amount of content being published by bots. Monitoring 37 trending local hashtags on Twitter from June 23-30, the report stated that, 'Propaganda-driven human-bots often used to push or engineer political campaigns, are deployed to harass, heckle and intimidate political rivals and journalists,' adding that some messages also incited direct violence."[9]

Twitter also stands accused of caving in to the Indian government pressure to remove pro-Kashmiri accounts critical of the state, according to Newsweek:

> "Twitter has been accused of bowing to Indian censorship and suppressing freedom of speech in Kashmir, after nearly one million tweets were removed.
>
> Almost 100 accounts were also made inaccessible to locals in the last two years, spurring claims that Twitter is contradicting the very values it purports to uphold. The findings were revealed in a study by the Committee to Protect Journalists (CPJ) on Friday, showing that Twitter agreed to block more accounts in the region than in every other country combined.[10]"

While platforms may struggle in some countries to find an appropriate balance between gaining permission to operate a service (which is valuable to activists), and avoiding threats of a ban or short term shutdown (which would harm activists), human rights groups, including Front Line Defenders, note that all too often platforms act too quickly to remove content or the accounts of legitimate protest groups, activists and journalists, without understanding the context for the posts, or investigating the legitimacy of the individuals and groups. Or, in the case of Apple, to bow to Chinese state pressure to remove a valued app activists in Hong Kong were using to safely coordinate protests to avoid riot police and teargas attacks[11].

Women activists and journalists are particularly likely to be the target of threats and physical attacks, even murders, that are initially encouraged in online threats. Pakistan offered a typical example in 2018:

> "Journalist Saba Eitizaz was forced to flee Pakistan after becoming the target of an online campaign that resulted in repeated online threats and sexual harassment, hacks, doxxing and character assassination. The campaign followed after her reporting on several high-profile stories of human rights abuses."[12]

When women activists are targeted, their children are also often included in threats. This growing tide of threats and violence against women activists and journalists was the focus of specific concern from the UN High Commissioner for Human Rights Zeid Ra'ad Al Hussein, who issued a statement in 2018 that included a litany of horrific examples. The statement noted:

---

[9] "Pakistan's Internet Landscape 2018" https://bytesforall.pk/sites/default/files/Internet%20Landscape%20Report%202018.pdf p.54

[10] Haddad, Tereq: "Twitter removes almost 1 million tweets, accused of bowing to Indian censorship" https://www.newsweek.com/twitter-removes-almost-1m-tweets-kashmir-accused-bowing-indian-censorship-1467721

[11] Hale, Erin: "Hong Kong protests: Apple pulls tracking app after China criticism" https://www.theguardian.com/world/2019/oct/10/hong-kong-protests-apple-pulls-tracking-app-after-china-criticism

[12] "Pakistan's Internet Landscape 2018" https://bytesforall.pk/sites/default/files/Internet%20Landscape%20Report%202018.pdf p.38

> “Online campaigns against women human rights defenders and organisations aim to damage their credibility as advocates, to diminish or obliterate the power of their voices, and to restrict the already limited public space in which women's activists can mobilise and make a difference.
>
> “The impact can be profound. The anxiety and fear suffered by the victims are compounded by a very real possibility of physical harm, as well as damage to livelihoods generated by the dissemination of false and sexually explicit images or other malicious lies. In a survey of eight countries last year, Amnesty International found at least 41 per cent of women who had been abused online feared for their physical safety, and 24 per cent feared for their family's safety, since online mobs who attack women often issue detailed and graphic threats against their children. Moreover, these attacks are extremely frequent and widespread. In 2014, the EU Fundamental Rights Agency found nearly a quarter of women surveyed had experienced online harassment.”[13]

_____

Clearly, this is a ferocious problem, and the challenges in meaningfully countering these threats and harms are enormous. But arguments to ban platforms, or for activists to join campaigns to leave the platforms, are meaninglessly privileged proposals from those unaware of the critical importance of the platforms to pro-democracy activists. It is essential to reiterate that platforms such as YouTube, Twitter and Facebook, and easy to use and widely-embraced encrypted messaging service such as WhatsApp, “are a lifeline for activists,” as Front Line Defender’s Andrew Anderson emphasises. Through such services, activists wage campaigns, communicate privately and safely, protect their friends and contacts, and document and share evidence of human rights violations.

However, the platforms need key reforms that will not only protect pro-democracy activists but help remedy abuses everywhere:

1) The platforms should work more closely with trusted regional and local NGOs to better understand the context behind and determine the validity of government requests for content and account takedowns. While social media platforms do liaise with some advisory groups in

---

[13] Statement by UN High Commissioner for Human Rights Zeid Ra’ad Al Hussein, 38th session of the Human Rights Council https://www.ohchr.org/EN/HRBodies/HRC/Pages/NewsDetail.aspx?NewsID=23238&LangID=E. The statement also says: “Several of this Council's own Rapporteurs have been the victim of online threats of violence and sexual violence; one, for example, was sent a graphic video of a person being decapitated. Mexican activists involved in sexual and reproductive rights have reportedly been targeted by death threats, harassment in public spaces, and acts of intimidation, including against their children. In Vietnam, following a series of online attacks, environmental activist Le My Hanh was physically attacked last year, with the video of the attack further disseminated on social media. In India, Gauri Lankesh, a journalist who published criticism of Hindu extremism, was killed last year following widespread online calls for violence against her; and her colleague Rana Ayyub has been subjected to thousands of hate-filled messages, including calls for her to be gang-raped and murdered, with dissemination of her phone number and home address.

In Italy, the speaker of Parliament, Laura Boldrini, has bravely faced down innumerable death threats and threats of sexual torture; the mayor of one town suggested on Facebook that a convicted rapist should be sent to her house "to put a smile on her face". In what has been termed GamerGate, thousands of anonymous online threats of murder and rape have targeted women who protested misogyny in video game culture. In Canada, a murderous rampage in April which killed 10 people and wounded 14 – most of them women – was reportedly motivated in part by the alleged perpetrator's radicalisation through so-called "incel" hate groups online. And in Iraq, several women candidates for parliament have reportedly faced online defamation campaigns, including the spread of faked photos and videos intended to intimidate and discredit them.”

some countries, this is sporadic. NGO reports cited in this statement all point to a lack of interaction with legitimate human rights bodies that are more knowledgable about their own countries.

2) The sheer scale of all these platforms creates an environment hopelessly open to abuse and exploitation. Platforms are inadequately staffed to monitor at a meaningful level. Algorithms cannot do the job because they are weak at recognising context, for example from keywords. For vulnerable human rights defenders, algorithmic decisions may determine freedom or imprisonment, even life or death. If the platforms need more human moderation at local scale, they should be required to provide it, and if necessary, change their business model, or be broken apart, to do so. Despite moving towards news curation and even creation, the platforms currently escape without the oversight obligations and responsibilities of media publishers. But they are not simply neutral, empty vessels for content. If platforms do not fit outdated, analogue definitions of a publisher, then a new hybrid category, complete with properly scaled duties and responsibilities, is needed in law.
3) The secretive and regularly evolving nature of social media and search algorithms makes them opaque to the users of the platforms, outside observers and researchers and therefore, it is hard to properly analyse the component problems or propose solutions. (See, for example, the recent analysis of online campaign advertising in recent Irish elections which noted the troubling paucity of data provided by Google, Facebook and Twitter.[14]). Regulators need to mandate algorithm transparency.
4) Platforms need to be more aware and vigorous and nuanced in assessing the possible future harms and interferences caused by their actions and programmes. The inadvertent support given the Duterte regime by Facebook provides a salutary case study.
5) Governments and regulators must foreground risks to activists as they consider ways to manage online problems. Too many proposals, such as banning anonymity, or placing deliberate back door vulnerabilities into encrypted messaging, undermine the grassroots movements towards democracy such nations claim to support. Such proposals are as much a threat to democracy as the online harms and electoral interference they hope to combat.
6) States and regulators will never address these broad problems unless they address the core business model of the platforms: micro-targeted advertisements based on data gathering at massive scale. As long as this model is allowed, driving in turn the prioritising of clickbait material that fuels engagements, and allowing ads and posts to be hidden from a larger audience that might refute their claims, the platforms (and we) will remain ripe for exploitation. Micro-targeting offers only minute benefits to the regular platform user — a more timely shoe or holiday ad. Set against this, pro-democracy activists are exposed to serious threats and violence, and all of society pays the price of the inevitable destruction of the norms of democracy.

Finally, as so many of these reports from activists and human rights organisations observe with dismay, democracy as a global aspiration is undermined when the political leaders of the world's notable democracies employ social media to lie, discredit, spread disinformation, and question and undermine election outcomes. Dictators and autocrats in the most repressive of nations are then emboldened and legitimised in doing the same[15]. Activists are disempowered — or worse. And long-standing democracies are gradually leached of any moral authority they once held. While platforms provide a vehicle for such individuals, their placement in positions of power is a sobering reminder that equally daunting societal issues are linked to the problems under consideration today by the Committees. Thank you.

---

[14] "As the researchers' state "…inconsistency across the three companies results in a systematic lack of transparency and comprehensive understanding of political and issue-based advertising online'. This outcome reflects the overall experience in the other European countries that participated in the ERGA monitoring process." "Elect Check 2019" https://fujomedia.eu/wp-content/uploads/2019/09/Elect-Check-2019-Report_Interactive-PDF-1.pdf p.6

[15] Filipina Maria Ressa spoke of the US having "a president so much like ours whose attacks against the press (and women) give permission to autocrats (like ours) to unleash the dark side of humanity and extend their already vast powers with impunity, especially in countries where institutions have crumbled." Bump, Philip: https://www.washingtonpost.com/politics/2018/11/21/warning-offered-by-filipina-journalist-targeted-by-her-countrys-president/

Biography:

Karlin Lillington is a technology columnist for the Irish Times with a special interest in privacy, data protection, information security, and digital and human rights. She also has written for a wide range of other publications including The Guardian, New Scientist, The London Sunday Times, Wired.com and Salon.com, and is a regular speaker and moderator at conferences and contributor on Irish radio and the BBC. She served as a non-executive director of Irish national broadcaster RTE and is currently on the advisory board of The Science Gallery at Trinity College Dublin and the board of the Dublin International Piano Competition. She has a PhD from Trinity College Dublin.

Appendix

Front Line Defenders Statement on harmful content and hate speech online

Front Line Defenders acknowledges the vital role that social media platforms play in the advocacy, networking and security strategies of human rights defenders around the world. However, the use and abuse of such platforms also presents numerous risks to defenders and the safeguards in place are not adequate in relation to the dangers faced.

Human rights defenders are targeted because of the vital work they do holding the powerful to account, exposing violations, defending the environment, promoting equality and seeking justice for victims. In response, through social media channels they are subjected to smear campaigns, threats, doxxing and hacking, often with a particular targeting of women activists.

These instances occur on a daily basis; examples include fake social media accounts being set up in the name of a prominent activist to discredit his work; over 500 violent comments left on a poem published by an LGBTI activist after she condemned a minister's anti-LGBTI tweet; and the personal details of over 200 human rights activists being released online and circulated via social media. In none of these cases did the relevant social media company respond quickly enough to shut down the accounts being used to attack the human rights defenders.

Years of evidence shows that when such attacks take place against human rights defenders, if no action is taken, they frequently lead to further, offline attacks. In 2017 Front Line Defenders analysed data on the killing of 312 activists. In 84 per cent of these cases where the necessary information was available, the defender had previously received a threat, many of which were made online. Smear campaigns which are given oxygen on social media serve the purpose in dozens of countries of delegitimising specific human rights defenders in the eyes of the public before they are arrested or attacked, reducing popular opposition to such moves.

What has been sorely lacking from parent companies thus far is both an understanding of the vital work human rights defenders do and an appreciation of the extremely risky — and varied — contexts in which they work, where comments which are permitted to accumulate targeting a specific person can and often do lead to offline violence, including murder. Front Line Defenders calls on companies operating social media platforms to develop institutional policies relating to human rights defenders taking into account the UN Declaration on Human Rights Defenders[16] as well as the EU Guidelines on Human Rights Defenders[17].

Setting up a specific mechanism to respond quickly to threats targeting human rights defenders is critical. This could be done in partnership with national and regional/international human rights organisations and should include context analysis of civil society space and rapid response mechanisms, where content targeting defenders in contexts which see a high level of attacks against human rights activists is prioritised, removed and bans imposed on the originators of such content.

---

[16] https://www.ohchr.org/en/issues/srhrdefenders/pages/declaration.aspx

[17] https://eeas.europa.eu/headquarters/headquarters-homepage_en/3958/EU%20Guidelines%20on%20Human%20Rights%20Defenders