GUIDELINES FOR MONITORING ONLINE VIOLENCE AGAINST FEMALE JOURNALISTS

Prof. Julie Posetti, Dr. Diana Maynard, and Nabeelah Shabbir

OSCE
the representative on freedom of the media
THERE IS NO SECURITY WITHOUT MEDIA FREEDOM

Guidelines for monitoring online violence against female journalists

Authors: Prof. Julie Posetti, Dr. Diana Maynard, and Nabeelah Shabbir
OSCE Project Manager: Paloma Madrigal Casanueva
Editorial Assistance: Goran Tanevski

Designer and artist: Imprimerie Centrale

Guidelines for monitoring online violence against female journalists

This document serves as a tool for monitoring and recording online violence against female journalists globally. It was commissioned by the OSCE Representative on Freedom of the Media and produced by researchers from the International Center for Journalists (ICFJ) and the University of Sheffield. It draws on research produced by the authors under commission to the UK's Foreign Commonwealth and Development Office (FCDO).

About the authors:

Prof. Julie Posetti (PhD) is VP of Global Research at the International Center for Journalists (US) and Professor of Journalism at City, University of London (UK). She is a multi award-winning internationally published journalist and academic with over three decades of experience, who has led several major UN-commissioned studies in the fields of disinformation, freedom of expression and the safety of journalists. She sits on the board (elect) of the International Fund for Public Interest Media (IFPIM), and she is a Senior Researcher affiliated with the Centre for Freedom of the Media and a Research Associate with the Reuters Institute for the Study of Journalism at the University of Oxford.

Dr. Diana Maynard is a Senior Research Fellow in the Computer Science department and a member of the Centre for Freedom of the Media, both at the University of Sheffield (UK). She has a PhD in Natural Language Processing (NLP), publishes extensively, and has more than 30 years of experience in the field. Her research involves multidisciplinary approaches to text and social media analysis, particularly concerning issues around the safety of journalists, online hate speech, and misinformation.

Nabeelah Shabbir is Senior Research Associate at the International Center for Journalists and a British-Pakistani freelance journalist. At The Guardian, she shared a British Journalism Prize with the 'Keep it in the Ground' team in 2015. She has co-authored a series of reports for UNESCO, and the Reuters Institute for the Study of Journalism at the University of Oxford. She has also worked at The Correspondent, the Financial Times, and Twitter.

This publication was commissioned in 2022 by the OSCE Representative on Freedom of the Media. It is part of the project "Safety of Female Journalists Online," or #SOFJO. The views, findings, interpretations, recommendations, and conclusions expressed herein are those of the authors and do not necessarily represent the official position of the OSCE and/or its participating States.

TABLE OF CONTENTS

Index

SECTION 1:
Introduction and context

Gender-based online violence aids and abets impunity for crimes against journalists. There is increasing evidence of a correlation, and even a causal relationship, between online threats towards female journalists and offline attacks.

According to research published by UNESCO and the International Center for Journalists (ICFJ),[1] **nearly three quarters (73%) of 714 international female journalists surveyed in late 2020 said that they had experienced online violence in the course of their work**. More alarmingly, **20% of them indicated that they had experienced offline attacks and abuse** that they believed had been seeded online.[2]

This suggests that gender-based online violence against journalists could be a predictor of physical violence, including murder with impunity. Online violence is also a feature of the enabling environment for the legal harassment and persecution of independent journalists.

It is therefore essential that online violence targeting female journalists be **effectively monitored, recorded and transparently reported** by the actors responsible for ensuring their safety - both online and offline - including the platforms, media employers, press freedom NGOs, and intergovernmental

1 See: 1) Posetti, J, Aboulez, N, Bontcheva, K, Harrison, J, and Waisbord, S, 2020: Online violence against women journalists: a global snapshot of incidence and impacts: https://unesdoc.unesco.org/ark:/48223/pf0000375136 (UNESCO) 2) Posetti, J, Shabbir, N, Bontcheva, K, Maynard, D and Aboulez, N, 2021: The Chilling: Global Trends in Online Violence against Women Journalists; Research Discussion Paper: https://unesdoc.unesco.org/ark:/48223/pf0000377223/PDF/377223eng.pdf.multi (UNESCO) 3) Posetti, J and Shabbir, N, 2022: The Chilling: A Global Study On Online Violence Against Women Journalists: https://www.icfj.org/our-work/chilling-global-study-online-violence-against-women-journalists (ICFJ-UNESCO)

2 op.cit.: Posetti et. al., 2020.

organizations such as the OSCE. Only then, can key responsive organizations and mechanisms take more effective protective action.

To that end, this tool has been produced to guide the monitoring and recording of online violations against female journalists to aid key responders in their efforts to prevent the escalation of online violence to offline harm. The tool presents a set of 15 research-derived indicators for online violence escalation (with examples of manifestations and tailored monitoring guidance), a gendered online violence typology, and examples of violations mapped to international codes and standards. The aim is to support a monitoring approach which **helps increase awareness of online violence threat escalation** along with a standardised approach for recording violations designed to systematise reporting.

Defining gender-based online violence

Online violence against female journalists is generally **sexist and misogynistic**. It frequently involves threats of physical and/or sexual violence; sexualised abuse and harassment; digital privacy and security breaches that can expose identifying information and exacerbate offline safety threats facing the target; and networked or mob harassment. It is also often bound up with **gendered disinformation** and its incidence and impacts are worse at the intersection of multiple forms of discrimination (e.g., racism, religious bigotry, homophobia).

The patterns of online violence against female journalists verge from large-scale attacks or extreme threats at a moment in time, through to the slow-burn of networked gaslighting,[3] which involves constant lower-level abuse.

Eight features of gender-based online violence[4]

Online violence targeting female journalists manifests itself in a variety of ways, but it has a number of common characteristics:

1. **It is generally misogynistic:** Misogyny is the dominant feature of online violence targeting female journalists.

3 See point 6. below.

4 Adapted from a taxonomy of online violence developed by the lead author for The Chilling, op.cit: Posetti, J. and Shabbir, N., 2022.

2. **It is frequently networked:** Online violence is often organized, coordinated or orchestrated. It can include State-sponsored 'sock puppet networks'[5] along with acts of 'patriotic trolling',[6] and involve mobs who seed hate campaigns within one online fringe network, before pushing it into more mainstream networks. But such abuse can come from individuals united in a common cause like misogyny.

3. **It radiates:** Perpetrators of online violence often target female journalists' families, sources, colleagues and supportive online communities, too.

4. **It is intimate:** In detail and delivery - the threats are personal. They arrive on mobile phone screens first thing in the morning and last thing at night, in private spaces as well as the newsroom, and they are often highly sexualised.

5. **It can be extreme, intense and prolific:** This often results in targets describing attacks in association with extreme weather events, natural disasters, and war such as: "torrential", "tsunami", "flood", "avalanche", "barrage", "trench warfare", "bombardment".

6. **It can behave like 'networked gaslighting':** constant moderate-low volume abuse and harassment that burns slowly but can be cumulatively devastating and undermine the target's confidence in her understanding of reality.

7. **It is more extreme in the context of intersectional discrimination** (e.g. race, sexual orientation, faith). These factors appear to attract increased exposure and worse impacts.

8. **It intersects with a threefold function of disinformation: a)** Reporting on disinformation and intertwined issues, such as digital conspiracy networks, conflicts, and far right extremism, is a trigger for heightened attacks on female journalists; **b)** disinformation tactics are routinely deployed in targeted multiplatform online attacks against female journalists; **c)** disinformation purveyors operationalise misogynistic abuse, harassment and threats against female journalists to undercut public trust in critical journalism and facts in general.

5 The term 'sock puppet' refers to a user account 'controlled by an individual (or puppetmaster) who controls at least one other user account', often for 'malicious and deceptive' purposes, and 'to manipulate public opinion': Kumar et al., 2017: https://arxiv.org/abs/1703.07355

6 Appropriation of notions of national loyalty in order to discredit other actors as "traitors".

Why is this tool needed?

The purpose of monitoring acts of harassment, intimidation and violence against journalists per se is to **help prevent the escalation of such attacks**, including to the level of offline harm. This approach is supported by the **UN Sustainable Development Goals (SDGs)**. SDG 16 seeks to: **'Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels'**.[7] Target 16.10 aims to: **'Ensure public access to information and protect fundamental freedoms, in accordance with national legislation and international agreements'**. Under this target, indicator 16.10.1 is concerned with the **'number of verified cases of killing, kidnapping, enforced disappearance, arbitrary detention and torture of journalists, associated media personnel, trade unionists and human rights advocates in the previous 12 months'**. This is supplemented by OSCE Ministerial Council Decision No. 3 on the Safety of Journalists from 2018,[8] in which the OSCE participating States recognized the devastating effects of online harassment on women journalists, acknowledged the threats to media freedom posed by attacks on female journalists in relation to their work, including through digital technologies. Participating States further committed to include a gender perspective in addressing the safety of journalists.

According to some estimates, approximately 50% of the SDG indicators identified for measuring and monitoring the achievement of targets **lack acceptable country coverage and agreed-upon methodologies**.[9] Related to this, other issues affecting monitoring include: access to data (treated proprietorially by social media companies); weak and inconsistent methodologies; perceived Western bias; and problems with comparisons especially concerning the circumstances of violations against journalists, longitudinal accuracy (since methodologies typically vary over time and per actor), and numerous issues surrounding data collection and categorisation, including not only the definition of violations but also with regard to who is (and who is not) included in counts.

Data beyond counts of killed female journalists is sparse and inconsistent, while data about online threats, attacks and abuse is even more patchy.

7 Sustainable Development Goal 16: https://sdgs.un.org/goals/goal16

8 Organization for Security and Co-operation in Europe (OSCE) Ministerial Council Decision No. 3 on the Safety of Journalists, 2018: https://www.osce.org/files/mcdec0003%20safety%20of%20journalists%20en.pdf

9 https://www.cgdev.org/blog/230-indicators-approved-sdg-agenda

Current inadequacies in the monitoring of online violence against female journalists can be broken down into **three primary issues: the lack of access to data, the lack of methodological consistency, and the lack of adequate monitoring.**[10]

But there are also significant gaps in the areas of gender-responsive monitoring and reporting on female journalists experiencing digitally-enabled attacks. This includes the need to account for the increased risks associated with networked misogyny and the essential requirement for gender-disaggregated approaches to data gathering and analysis. There is, therefore, a clear need for more systematic monitoring and reporting of online violations against female journalists, adopting a human rights-based approach.[11]

Key monitoring gaps:

- The need to deploy **gender-sensitive** and **gender-responsive** approaches to data collection and distribution
- The need for **digital literacy** and **digital data analysis** skills within organizations that undertake monitoring
- The need for **guaranteed and free access to large datasets** held by social media companies for the purpose of monitoring and escalating responses to safety threats experienced by female journalists online

10 Harrison, J., Maynard, D. and Torsner, S. Strengthening the Monitoring of SDG 16.10.1 and the Manifestations of Violations against Journalists through an Events-based Methodology. Journal of Media and Communication, vol. 8, no. 1: Rethinking the Safety of Journalists, 2020.

11 ibid.

The need for more sophisticated contextual analysis

Online attacks and threats can take many forms as described above, and **it is important that the whole spectrum of abuse is monitored rather than just the most extreme instances**. While relatively 'minor' abuse may seem trivial in isolation, when delivered at scale, over a long period of time, and as part of networked and/or organized 'pile-ons', its effect can have severe psychological consequences and can indicate potential escalation to serious harm. Similarly, since online violence radiates and this radiation can in and of itself be an indicator of escalation, attacks on the targeted journalists' family members, colleagues and sources are significant, and they should be recorded and monitored too.

Preventing online violence against female journalists requires understanding the complex nature and entire temporal and situational context within which the abuse takes place. Rather than simply relying on counts of incidents, this entails systematic studies of online messages (including those contained in audio, video and image-based content) and the analysis of abuse within the wider context of messages to and from a journalist under attack. **It also necessitates methods, tools and indicators which can then be developed to detect, predict and ultimately help prevent the escalation of online abuse, harassment and attacks against female journalists, into even more serious situations - both online and offline.**[12]

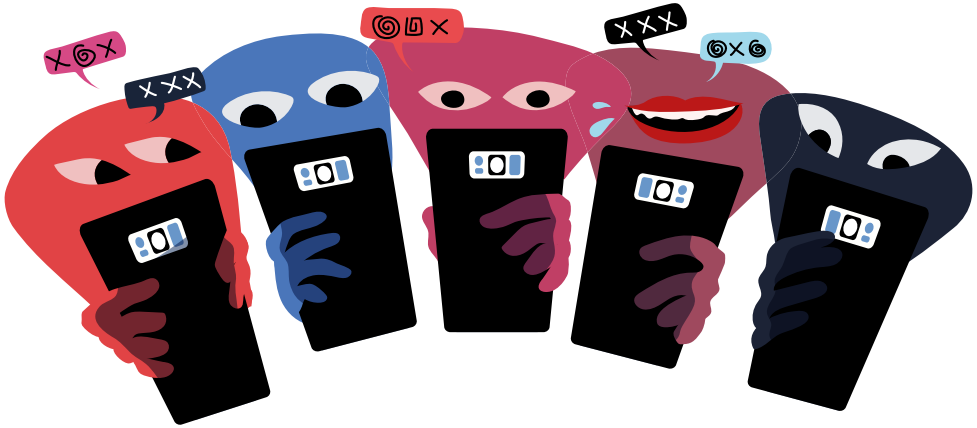
This approach demands **more advanced digital literacy and digital data analysis skills** within organizations undertaking such monitoring, from news outlets and civil society organizations, through to States and intergovernmental actors. While digital literacy skills can be developed through training and knowledge sharing among individual and organizational actors, in organizations conducting monitoring where digital data gathering and analysis capability does not exist and cannot be developed, collaboration with experts (e.g. computer scientists) will likely be necessary.

12 The authors are currently engaged by the UK's Foreign Commonwealth and Development Office to develop a research-led Online Violence Alert and Response System.

Data access challenges

Effective monitoring, recording and reporting of online violence escalation affecting female journalists depends on free and guaranteed **access to large datasets held by social media companies** (e.g., Meta, Google, Twitter, Tiktok) for independent researchers, and expert actors within civil society and journalism. However, these companies increasingly restrict and monetise such access to the detriment of these organizations' efforts to secure female journalists' safety. Simultaneously, they are reducing their human rights policy and safety staff, undermining their expertise and capacity to internally monitor violations.

This is why it is **vitally important that regulatory efforts focused on social media include requirements for transparency and accountability**, starting with direct access to data pipelines for researchers within academia and civil society organizations, along with news organizations committed to ethical and privacy preserving data collection and analysis practice. Tech companies should also be required to monitor and report transparently on online violence towards female journalists (and other high groups) in every language and region in which they operate.



SECTION 2:

15 key indicators for online violence escalation

The following **15 indicators for online violence escalation** can be used by social media companies, intergovernmental organizations, States, news outlets, civil society and academia to inform threat assessments and guide the monitoring and recording (see Section 3) of online violence against female journalists.

These research-derived indicators reflect the signals associated with evolving monitoring and response systems designed to trigger intervention in cases of female journalists under attack online.

They draw on:

- **9 interdisciplinary** Big Data Case Studies based on regionally and globally emblematic online violence cases to help inform OSCE responses: Carmen Aristegui (Mexico), Rana Ayyub (India), Carole Cadwalladr (UK), Daphne Caruana Galizia (Malta), Ghada Oueiss (Lebanon), Maria Ressa (the Philippines), Marianna Spring (UK), Ferial Haffajee and Pauli van Wyk (South Africa)[13]
- **Contextual research** in 15 countries[14]

13 These Big Data Case Studies, involving analysis of approximately 18 million social media posts, were produced by ICFJ researchers and University of Sheffield computer scientists under commission to UNESCO and the UK Foreign, Commonwealth and Development office (2021-2023). They feed an Online Violence Alert and Response System currently being developed by the authors: https://www.icfj.org/our-work/online-violence-big-data-case-studies

14 Brazil, Mexico, Poland, Serbia, the UK, the US, Sweden, Sri Lanka, Pakistan, the Philippines, Kenya, Nigeria, South Africa, Lebanon, Tunisia. Produced by ICFJ for UNESCO by an international team of researchers under commission to UNESCO. See the country case study summaries here: https://www.icfj.org/our-work/chilling-global-study-online-violence-against-women-journalists (pp 250-311).

- **Expert consultations** with international journalists, human rights lawyers, civil society organizations and subject matter experts[15]
- Validation via a series of **OSCE RFOM-led consultations** with journalists, editors, civil society organizations and Member States' representatives[16]

Each Indicator is accompanied by **tailored monitoring guidance for responders**. A number of features of the monitoring guidance are duplicated across the Indicators to allow for an assessment of an isolated incident being recorded in detail against an indicator.

15 Indicators for gendered online violence escalation

1. Death and rape threats

Rape and death threats are features of targeted online violence against female journalists which indicate serious propensity for escalation. 25% of women-identifying journalists who responded to the ICFJ-UNESCO online violence survey (2020) reported **receiving death threats or threats of general physical violence**, while 18% said they **had experienced threats of sexual violence**. Every one of the emblematic cases studied by the authors had received multiple rape and death threats online. **One of them - Daphne Caruana Galizia - was assassinated in Malta in 2017 after being targeted in an online violence campaign that included such threats.** Notably, the Public Inquiry into the murder of Caruana Galizia pointed to the **connection between the online violence directed at the journalist and her brutal killing.**[17] While a single credible threat of death or rape should be treated very seriously and reported to the relevant authorities where appropriate, when such threats are received at scale, or when they arrive in close succession, the risk is extreme.

Other examples from the OSCE region include:

- **Sevgil Musaieva** (Ukraine) received a Facebook message in June 2022 that read: "Sevgil, I have a feeling that your throat is going to be cut. I don't know who and when – but run." She and the staff of the online outlet she edits, Ukrainska Pravda, received (mostly anonymous) death threats for reporting on local government officials, posted publicly online and via personal messages sent to their social media accounts and phones.[18] [19]
- **Jelena Obućina** (Serbia), a journalist from Nova S TV, received direct messages on Twitter threatening her with impalement and being burned, in a sexualized context.[20]
- **Hale Gönültaş** (Turkiye), a prominent reporter for online news portal Gazeteduvar, recalled one threatening tweet: "May you drown in your own blood". In response to her reporting on jihadist armed group ISIS activities in Turkiye, she also received an email with a video of a boy beheading an unidentified male. For the next few days, she received phone calls originating from Raqqa, a Syrian town under ISIS rule at the time. A few days after publishing an investigative piece on thousands of missing Yazidi women, abducted by armed jihadist groups in Iraq, Gönültaş received a phone call from a man who clearly and calmly threatened her (in Turkish) with death and repeated her address. She was temporarily assigned a remote protection officer by Turkish authorities as a result.[21]
- **Arzu Geybullayeva** (Azerbaijan) first received death threats via Facebook in 2014, where she was a reporter for *Agos*, a Turkish Armenian newspaper. Geybullayeva was told how she should be killed and where she should be buried. More rape and other sexual threats followed. These threats continued in 2020, after she relocated for her own safety.[22]
- Serbia's **Jovana Gligorijević** has written about being threatened sexually.[23] She continues to be targeted for speaking out about the abuse she experiences, with online 'jokes' from trolls suggesting she should kill herself. In 2020, during the COVID-19 pandemic, she also

15 Including three research consultations with over 30 international experts and practitioners in Perugia and New York 2022-2023.

16 Including events in Skopje, Bishkek and Vienna 2022-2023.

17 Public Inquiry, Daphne Caruana Galizia Foundation, 2021: https://www.daphne.foundation/en/justice/public-inquiry

18 Ukrainian journalists Sevgil Musaieva and Sonia Lukashova receive death threats, CPJ, 2022: https://cpj.org/2022/06/ukrainian-journalists-sevgil-musaieva-and-sonia-lukashova-receive-death-threats/

19 Ukraine: Online Harassment Of Journalists At Ukrainska Pravda Following Report On Politicians – CFWIJ Calls On Ukrainian Authorities To Investigate, CFWIJ, 2023: https://www.womeninjournalism.org/threats-all/ukraine-online-harassment-of-journalists-at-ukrainska-pravda-following-report-on-politicians-cfwij-calls-on-ukrainian-authorities-to-investigate

20 Terrible Threats to Journalist Jelena Obućina in Serbia, Safejournalists.net, 2022: https://safejournalists.net/portfolios/safejournalists-terrible-threats-to-journalist-jelena-obucina-in-serbia/

21 Turkish journalist receives threat after story on sale of Yazidi girl by ISIL in Ankara, Stockholm Center for Freedom (SCF), 2018: https://stockholmcf.org/turkish-journalist-receives-threat-after-story-on-sale-of-yazidi-girl-by-isil-in-ankara/

22 Azerbaijan: Hateful Online Attacks Against Journalist Arzu Geybullayeva Over Her Views, CFWIJ, 2020: https://www.womeninjournalism.org/threats-all/azerbaijan-hateful-online-attacks-against-journalist-arzu-geybullayeva-over-her-views and via OSCE #SOFJO, 2019: https://www.osce.org/files/f/documents/5/9/391031.pdf

23 op.cit.: Posetti and Shabbir, 2022 (p80).

received an unambiguous death threat on her Instagram account. The perpetrator was arrested shortly after she reported the threat.

- Brussels-based **Tanja Milevska**, a freelance journalist from North Macedonia, is threatened online when she writes about geopolitical issues. In 2020 she tweeted that she'd received the following message: "If the lights go out in your entrance hall, I advise you to get on your knees and pray."[24]
- Finnish investigative journalist for national broadcaster Yle, **Jessikka Aro**, received a late-night phone call from a number in Ukraine playing gunfire. She got text messages and emails calling her a "NATO whore", "NATO drug dealer". A message from her dead father claimed to be "watching her."[25]
- **Rena Netjes** (Netherlands) said she received death threats which came in quick succession in January 2023 from a Kurdish rebel group on the Turkish Syrian border.[26]
- **Boroka Paraszka** (Hungary/ Romania), who works for a Hungarian-language public radio station in Romania received death threats related to her reporting on human rights and minority rights. In November 2022, a politician from the Hungarian far-right Our Homeland Movement spoke publicly about hanging or 'eliminating' the journalist.[27]
- Macedonian journalist **Meri Jordanovska** has experienced online violence since 2009.[28] When she publishes critical reporting on politics, she said she typically receives rape threats on her Facebook account.
- In 2014, a **Swedish journalist** reported a series of threats she received to the police.[29] However, the court found that the following threatening statement was protected by freedom of speech provisions, due to its general nature: "To me gender equality is when you take a sexist feminist whore in the vagina with a large knife".
- **Silvia Bencivelli** (Italy), who was subjected to incitement to rape via a YouTube video in retaliation for her counter disinformation reporting.[30]
- **Natalia Żaba** (Poland) received non-stop harassment from a perpetrator sending pornographic images and describing sexual situations he imagined while she was reporting from the Balkans.[31]
- **Joanne Chiu** (Canada), a senior reporter at Toronto Star: "They said I should get my neck ready because they were coming over to my house to behead me - horrifying".[32]
- **Apoorva Mandavilli** (US) at The New York Times received a similar death threat:[33] "The emails are actually worse, because they're more private... that I should have my head cut off in public...I am a liar. I should be ashamed of myself. I don't deserve to live. One said, I hope you get the virus and choke...very nasty and vile emails".
- **Sharon Ní Bheoláin** (Ireland)'s case is 'an early example of a harasser producing what are now known as deepfakes'.[34] The RTÉ journalist and presenter's perpetrator was sentenced to four and a half years in prison in 2018 for harassing her. The officers also uncovered 217 private messages in which he named Ní Bheoláin while discussing torture, murder and extreme sexual violence.
- **Tímea Karip** (Hungary) of Index.Hu said in 2016 that she received "hardcore porn images via email along with comments describing her participation in forced sexual intercourse" which was partly "why some female journalists intentionally left their bylines off particularly sensitive articles and disguised their Facebook identities - politics and being a woman are both risk factors" for harassment".[35]

24 #ReportIt: Sharing solidarity with journalists facing sexist online harassment, Mapping Media Freedom, 2021: https://www.mappingmediafreedom.org/2021/03/08/reportit-sharing-solidarity-with-journalists-facing-sexist-online-harassment/ and Journalist Tanja Milevska Gets Hate Speech And Sexist Comments Over Her Coverage On The Election Campaign, CFWIJ, 2020: https://www.womeninjournalism.org/threats-all/northmacedonia-journalist-tanja-milevska-gets-hate-speech-and-sexist-comments-over-her-coverage-on-the-election-campaign

25 Effort to Expose Russia's 'Troll Army' Draws Vicious Retaliation, NYT, 2016: https://www.nytimes.com/2016/05/31/world/europe/russia-finland-nato-trolls.html?_r=0 and Jessikka Aro: How pro-Russian trolls tried to destroy me, BBC, 2017: https://www.bbc.com/news/blogs-trending-41499789

26 Netjes in Villamedia, 2023: https://www.villamedia.nl/artikel/rena-netjes-ontvangt-expliciete-doodsbedreigingen-wat-ze-willen-is-dat-ik-stop-met-schrijven-dat-ga-ik-niet-doen

27 Hungarian Politician Discusses the Hanging of Journalist Boróka Parászka, Council of Europe Safety of Journalists Platform, 2022: https://fom.coe.int/en/alerte/detail/107638184;globalSearch=false

28 Female Journalists Attacked Online | Meri Jordanovska, Birn Balkans YouTube page, 2019: https://www.youtube.com/watch?v=kGOn2aln09s

29 op.cit.: Posetti and Shabbir, 2022 (p218).

30 How to deal with trolls, conspiracy theorists and hoax spreaders on the web, IJF, 2014: https://www.journalismfestival.com/news/how-to-deal-with-trolls-conspiracy-theorists-and-hoax-spreaders-on-the-web/

31 As told to the authors, op.cit.: Posetti and Shabbir, 2022.

32 I've lost track of how many threats I've received. That's how common online hate is, Toronto Star, 2021: https://www.thestar.com/news/canada/2021/06/29/ive-lost-track-of-how-many-threats-ive-received-thats-how-common-online-hate-is.html

33 op.cit.: Posetti & Shabbir, 2022 (p150).

34 Legal Responses to Online Harassment and Abuse of Journalists: Perspectives from Finland, France and Ireland, OSCE and IPI, 2021: https://www.osce.org/files/f/documents/1/6/413552.pdf

35 At Hungary's Index.hu, online abuse 'comes with the job', IPI Media, Re:Baltica, 2016: https://ipi.media/at-hungarys-index-hu-online-abuse-comes-with-the-job/

- **Vilja Kiisler** (Estonia), was targeted in 2019 on Facebook by a troll who had also harassed politicians: "Your judgement day will soon arrive where the boomerang you threw will come back to you with great punishment."[36] She combined numerous threats received in one complaint and went to the police, but 'nothing happened.'
- **María Morán** (Spain) was targeted with rape threats and her 18-month-old was insulted (i.e. called a bastard child, not recognised by a footballer father), after the sports journalist asked a question in a football press conference in May 2023.[37]
- **Ada Borowicz** (Poland) was criticized and threatened with rape online after 'omitting' to report that the perpetrators of a crime were migrants: "This reporter should be raped."[38]

Monitoring guidance for responders

- **Record the threat/s** (describe the threat, attach a screen grab of the threat, include the URL where relevant, include evidence of any image based abuse featured e.g., photos, memes).
- **Identify the perpetrator/s** (e.g., username/handle; location [using geolocation information or other forensic tools]; mobile number used; email address; real name and affiliations where evident).
- **Identify the medium/vectors/facilitators of the threat/s** (e.g., social media, chat app, text message, email).
- **Identify and monitor any hashtags** being used in connection with the abuse.
- **Conduct a risk-assessment for the perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State or foreign State actor, the risk is heightened).

36 "You will collect your teeth with broken fingers". Why haters are getting away with online abuse, Re:Baltica, 2021: https://en.rebaltica.lv/2021/09/you-will-collect-your-teeth-with-broken-fingers-why-haters-are-getting-away-with-online-abuse/

37 La periodista María Morán denuncia graves amenazas por una pregunta a Ancelotti: "Que me violen, que mi hija es bastarda...", 20 Minutos, 2023: https://www.20minutos.es/deportes/noticia/5123859/0/periodista-deportiva-maria-moran-amenazas-pregunta-ancelotti/?utm_source=twitter.com&utm_medium=socialshare&utm_campaign=mobile_web

38 Targeting the messenger: Journalists face an onslaught of online harassment, Index on Censorship, 2019: https://www.indexoncensorship.org/targeting-the-messenger-journalists-face-an-onslaught-of-online-harassment/

- In serious cases, where the threats have multiplied or where a high risk perpetrator is involved, **conduct a network analysis** to determine the original source of the threat, and map its distribution if the capability exists.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers, and to measure the speed and spread of attacks and pile-ons aimed at the journalist.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant abuse terms and tropes featured in abuse directed at the target to help determine the level of risk.
- **Determine if the threat has migrated across platforms** and monitor its spread, recording all violations.
- **Indicate if the threat has been reported to law enforcement.** Monitor the progress of the investigation/s.
- **Indicate if the threat has been reported to the company/platform facilitating the threat.** Monitor the progress of these investigations.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform,[39] UN Special Procedures, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists)[40] and monitor the follow-up.

2. *Identifiable or suspected State/foreign State actor, or political extremist involvement*

The risks posed in cases of targeted online violence against female journalists with **established or suspected State actor or foreign State actor involvement are extremely high**. We see evidence of this in the case of Al Jazeera principal Arabic presenter Ghada Oueiss who has been targeted in chilling online attacks by high ranking officials in both Saudi Arabia and

39 The Council of Europe, Reporters Without Borders, the International Federation of journalists, the European Federation of journalists, the Association of European journalists and ARTICLE 19 signed a memorandum of understanding in 2014; now 14 international NGOs and associations of journalists are partners to the platform, posting alerts (and sometimes, responses): https://www.coe.int/en/web/civil-society/platform-for-the-safety-of-journalists#:~:text=The%20Platform%20for%20the%20safety,Council%20of%20Europe%20member%20states

40 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

the UAE.[41] Similarly, dozens of BBC Persian service journalists have been targeted online by Iranian authorities representing an extreme threat of extraterritorial harm.[42] State-actor linked cases, such as **Maria Ressa** in the Philippines, are also highly relevant to the issue of impunity. Additionally, they frequently involve the operationalisation of partisan or State media to **instigate or fuel targeted online attacks** on female journalists (see Indicator 15). **This type of threat is heightened in the context of armed conflict.**

Examples from the OSCE region include:

- **Sevgil Musaieva** (Ukraine), who described "psychological attacks" coming from representatives of the subject of her outlet's reporting - wealthy Ukrainians in Dubai during war-time. Before the investigation was published, public relations specialists and political analysts said they believed an implicated politician was preparing to use a bot farm, or a network of manipulated social media accounts, for "an attack" on the journalists and Ukrainska Pravda'.[43]
- **Allison Morris** (Northern Ireland) who sued a politician for instigating pile-ons against her (and won) in 2019. She told the press outside the court, that she hoped the case sent "...a very strong message that women in the media, or in any other public role, are not open season for online abuse of a misogynistic nature".[44]
- Dutch journalist **Rena Netjes**, who said she received death threats in January 2023 from the Kurdish rebel group PKK and Syrian Kurdish militia called the People's Protection Units (YPG), which also involved hacking attempts.[45]
- **Hale Gönültaş** (Turkiye), Gazeteduvar, was repeatedly threatened with death after reporting on ISIS activities in Türkiye (See detailed example under Indicator 1.)[46]

41 1) Posetti, J., Maynard, D., al-Kaisy, A., Harb, Z., and Shabbir, N, ICFJ, 2023: https://www.icfj.org/sites/default/files/2023-02/ICFJ_BigData_Ghada%20Oueiss_Online%20Violence.pdf 2) also published as part of Forbidden Stories' Story Killers project: https://www.icfj.org/our-work/online-violence-big-data-case-studies

42 UN Raises 'Grave Concern' with Iran Over Harassment of BBC News Persian Staff, Doughty Street Chambers, 2022: https://www.doughtystreet.co.uk/news/un-raises-grave-concern-iran-over-harassment-bbc-news-persian-staff

43 1) op.cit.: CFWIJ, 2023: https://www.womeninjournalism.org/threats-all/ukraine-online-harassment-of-journalists-at-ukrainska-pravda-following-report-on-politicians-cfwij-calls-on-ukrainian-authorities-to-investigate and 2) Ukrainska Pravda journalists receive menacing messages online after report on politician, CPJ, 2023: https://cpj.org/2023/02/ukrainska-pravda-journalists-receive-menacing-messages-online-after-report-on-politician/

44 Ex-DUP MLA McCausland apologises to Irish News reporter, BBC, 2019: https://www.bbc.com/news/uk-northern-ireland-48258135

45 op. cit.: Villamedia, 2023.

46 op. cit.: SCF, 2018.

- **Evgenija Carl** (Slovenia) was called a "prostitute" on Twitter by the leader of the opposition party Janez Janša. Carl worked for the national Slovenian Television station (RTVSLO) at the time.[47]

Monitoring guidance for responders

- **Record the threat/s** (describe the threat, attach a screen grab of the threat, include the URL where relevant, include evidence of any image based abuse featured e.g., photos, memes).
- **Identify the perpetrator/s** (e.g., username/handle, location [using geolocation information or other forensic tools]; mobile number used; email address; real name and affiliations where evident).
- **Identify the medium/vector/facilitators of the threat/s** (e.g., social media, chat app, text message, email).
- **Identify and monitor any hashtags** being used in connection with the abuse.
- **Conduct a risk-assessment of the perpetrator** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are a known associate of a State actor or foreign State actor, the risk is heightened).
- **Monitor the perpetrator's online activities** to pre-empt potential escalation.
- In serious cases, where the threat has multiplied or where a high risk perpetrator is involved, **conduct a network analysis** to determine the original source of the threat, and map its distribution if the capability exists.
- **Deploy network analysis, and abuse monitoring and visualisation tools** in order to understand connections between abusers, and to measure the speed and spread of attacks and pile-ons aimed at the journalist.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant abuse terms and tropes featured in abuse directed at the target to help determine the level of risk.
- **Determine if the threat has migrated across platforms** and monitor its spread.

47 Called a prostitute by the prime minister, a Slovenian journalist tells her story, Global Voices, 2021: https://globalvoices.org/2021/03/09/called-a-prostitute-by-the-prime-minister-a-slovenian-journalist-tells-her-story/

- **Indicate if the threat has been reported to the company/platform facilitating the threat and/or law enforcement** and monitor the progress of these investigations.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform,[48] UN Special Procedures, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists)[49] and monitor the follow-up.

3. *Proximity to attackers and relative threat level associated with perpetrators (e.g, Presidents, organized crime gangs & paramilitaries)*

When online violence is perpetrated by powerful and/or dangerous entities (e.g., State actors, political leaders, religious leaders, government officials, criminal gangs, military and paramilitary operatives) the risk of escalation (including offline) increases significantly. **The more powerful or influential the abuser, the greater the risk of digital mobs piling on and physical mobs taking the violence offline.** In the ICFJ-UNESCO survey, 37% of the female journalists who responded identified political actors as primary perpetrators of the online violence they endured. **Extremist political figures** (e.g., far right political leaders) **represent a common threat.** The researchers recorded cases in the US, the UK, South Africa, Mexico, Pakistan, the Philippines and Brazil (among others) where senior political figures, including presidents, had instigated or fuelled online violence against female journalists. The proximity of the abuser to the target is also a relevant cross-cutting indicator - especially if they are within easy reach.

48 The Council of Europe, Reporters Without Borders, the International Federation of Journalists, the European Federation of Journalists, the Association of European Journalists and ARTICLE 19 signed a memorandum of understanding in 2014; now 14 international NGOs and associations of journalists are partners to the platform, posting alerts (and sometimes, responses): https://www.coe.int/en/web/civil-society/platform-for-the-safety-of-journalists#:~:text=The%20Platform%20for%20the%20safety,Council%20of%20Europe%20member%20states

49 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

Examples from the OSCE region include:

- Serbian journalist **Brankica Stanković**, who is the complainant in a number of ongoing cases regarding targeted online violence in Serbia.[50] She has reported on the ties between gangs and political figures. The police discovered that the barrage of on- and offline threats against her came from a group of the most notorious criminals in Serbia, who assassinated the first democratically elected Serbian Prime Minister, Zoran Đinđić, in 2003. Stanković previously lived under police protection for five years between 2004 and 2009.
- **Boroka Paraszka** (Hungary/ Romania), a Hungarian-language public radio station reporter in Romania, who has received death threats from a political figure (see Indicator 1 for further detail).[51]
- **Jovana Gligorijević** (Serbia), who was the subject of a 28-minute video by a YouTube influencer and two representatives of a far-right political organization who accused her of being "the main source of Serbia's downfall".[52]
- **Anthi Pazianou** (Greece), who was targeted in September 2017 after the leader of an extreme right-wing movement accused her of bias, posting a photo of her with a refugee football team which she had written about on the Greek island of Lesbos. It "sparked a wave of insults and verbal sexual harassment, both online and in the streets."[53] (See Indicator 8 for examples of offline harm connected to online violence).
- **Emilia Şercan** (Romania) reported on a plagiarism scandal involving a politician which triggered online attacks.[54] In this context, a Facebook message from an unknown person was sent to her[55] containing personal photos taken 20 years earlier by her then fiancé.
- **Ida Erämaa** (Finland), who was targeted by far-right politicians in 2023 and abused online with reference to her dating history.[56]

50 op.cit.: Posetti and Shabbir, 2022 (p66).

51 op. cit.: Council of Europe Safety of Journalists Platform, 2022.

52 op. cit.: Posetti and Shabbir, 2022 (p44).

53 Online harassment brings special risks for freelance journalists, IPI, 2017: https://ipi.media/online-harassment-brings-special-risks-for-freelance-journalists/

54 Romanian education minister resigns after plagiarism accusations, Reuters, 2022: https://www.reuters.com/world/europe/romanian-education-minister-resigns-after-plagiarism-accusations-2022-09-30/

55 Romanian Journalist Emilia Şercan Victim of Smear Campaign, Council of Europe Safety of Journalists platform: https://fom.coe.int/en/alerte/detail/107637394;globalSearch=true

56 Online threats against Finnish journalist Ida Erämaa should be investigated and condemned, Council of Europe, 2023: https://www.coe.int/en/web/commissioner/-/online-threats-against-finnish-journalist-ida-er%C3%A4maa-should-be-investigated-and-condemned

- **Patricia Devlin** from Northern Ireland was able to prove that she was being targeted by the criminal actors she was reporting on as a crime journalist for *Sunday World*. She linked Facebook groups where she was being targeted to credible death threats and visits from the police telling her not to report from an area for her safety, as well as graffiti on a wall of Belfast implying she should be killed.[57]
- **Nektaria Stamouli** (Greece) suffered online violence in 2022 after reporting on eroding press freedom standards in her country connected to a surveillance scandal implicating Greek intelligence services. The abuse was fanned when she was discredited by a government minister for her reporting.[58]

General monitoring guidance

- **Record the threat/s** (describe the type of threat [e.g., rape threat, death threat; threat of other physical violence; threat to harm others], attach a screen grab of evidence of the threat if relevant)
- **Identify the perpetrator/s** (e.g., location [using geolocation information and other forensic techniques], username/handle, include any photographic or video evidence of the abuser if available, along with real name and affiliations where evident).
- **Conduct a risk-assessment for the suspected perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor or political leader, the risk is heightened).
- **Identify the threat medium/vector/facilitators** (e.g., social media, chat app, email, text message).
- **Identify and monitor any hashtags** being used in connection with the abuse.
- **Monitor targeted abuse across all the platforms** where the journalist is present to help respond to pile-ons which extend the risk of offline harm.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers, and to measure the speed and spread of attacks and pile-ons aimed at the journalist.

57 op.cit.: Posetti and Shabbir, 2022 (p188).

58 Government spokesperson discredits Politico Europe correspondent Nektaria Stamouli, European Centre for Press and Media Freedom (ECPMF), 2022: https://www.mapmf.org/alert/25130

- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant abuse terms and tropes featured in abuse directed at the target to help determine the level of risk.
- **Indicate if the threat has been reported to law enforcement agencies and/or the company/platform** facilitating the threat, and monitor progress of these responses.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[59] UN Special Procedures) and monitor follow-up.

4. Threats associated with impunity cases

Perpetrators of online violence against female journalists increase risk levels by **threatening their targets with reference to cases of other journalists murdered with impunity**. For example, in Malta, **Caroline Muscat**, the Editor-in-Chief of The Shift is threatened with being killed in a bomb blast like her assassinated former colleague **Daphne Caruana Galizia**. In Northern Ireland, **Patricia Devlin** and staff of the *Sunday World* were threatened with a reminder of the historic killing with impunity of the paper's **Marty O'Hagan**. In India, **Rana Ayyub** is threatened online with meeting the same fate as her murdered friend **Gauri Lankesh** (who was subjected to online violence prior to her murder in 2017), while Lebanese journalist **Ghada Oueiss** is threatened with being murdered like her friend **Jamal Khashoggi**, who was assassinated inside Saudi Arabia's Istanbul consulate in 2018. And in Brazil, journalist **Talita Fernandes** of *Folha da São Paolo* was threatened with images of slain journalist **Vladimir Herzog**.

Further examples from the OSCE region include:

- **Female journalists with the BBC Persian service** (UK), who have been targeted with reference to Ruhollah Zam, the exiled journalist executed by Iran in 2020 after being lured to Iraq.[60]
- **Marta Jančkárová** (Slovakia), host of a political programme on public radio RTVS, was targeted via email and phone calls in February 2023

59 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

60 op. cit.: Posetti and Shabbir., 2022 (p90).

extreme right-wing opposition politician on her show had showed up instead of the scheduled politician for the programme, and when they were not let in, they held a press conference outside the radio and targeted the station's editors). After the death of Ján Kuciak, a journalist murdered alongside his partner Martina Kušnírová in Slovakia, the Investigative Centre of Ján Kuciak surveyed 400 journalists and found that two-thirds of them had experienced some form of threat of attack within the past year.[61] Online harassment was the most common threat and it often came via politicians with reference to Kuciak. The former prime minister Robert Fico had called journalists "dirty, anti-Slovak prostitutes" in 2016.

- **Leona O'Neill** (Northern Ireland) was abused after being at the scene when journalist Lyra McKee was killed by paramilitaries with impunity in 2019. A blogger accused her of McKee's death. She told CPJ in 2019: "The day after I escaped death in a shooting and had witnessed a colleague being murdered, I was faced with hundreds of messages calling for me to be attacked, stabbed, arrested, set on fire, that my children would burn in Hell, that I was a liar, and that I made up what happened for personal gain. This went on for months. I had to contact the police about several people who seemed to be obsessed with causing me harm and had solicited donations to "wage war" on me. Some messages tell me I am not welcome in certain areas of my city."[62]

Monitoring guidance for responders

- **Record the threat/s** (describe the threat, attach a screen grab of the threat, include the URL where relevant, include evidence of any image based abuse featured e.g., memes).
- **Identify the medium of the threat/s** (e.g., social media, chat app, text message, email).

61 1) Akcia Francesco: Vyhrážal sa redaktorke Jančkárovej, NAKA ho zadržala. Vlastnil zbraň, sudca však domovú prehliadku nepovolil, Aktuality.sk, 2023: https://www.aktuality.sk/clanok/BeDXKeC/akcia-francesco-vyhrazal-sa-redaktorke-janckarovej-naka-ho-zadrzala-vlastnil-zbran-sudca-vsak-domovu-prehliadku-nepovolil/ 2) Research: Do Slovak journalists feel safe? Ján Kuciak Investigative Center, 2023: https://icjk-sk.translate.goog/229/Vyskum-Citia-sa-slovenski-novinari-bezpecne?_x_tr_sl=sk&_x_tr_tl=en&_x_tr_hl=sk&_x_tr_pto=wapp 3) Journalism in Slovakia is under threat - on a daily basis, VSquare, 2023: https://vsquare.org/journalism-in-slovakia-is-under-threat-on-a-daily-basis/

62 Q&A: Leona O'Neill on the aftermath of Lyra McKee's killing in Northern Ireland, CPJ, 2019: https://cpj.org/2019/12/leona-oneill-aftermath-lyra-mckee-killing-northern-ireland/

- **Identify the perpetrator/s** (e.g., username/handle, location [using geolocation information and other forensic techniques], mobile number used, email address, real name and affiliations where evident).
- **Describe the impunity case referenced** and the potential implications for the journalist targeted online.
- **Monitor the perpetrator's online activities** to pre-empt potential escalation.
- **Identify and monitor any hashtags** being used in connection with the abuse.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers, and to measure the speed and spread of attacks and pile-ons aimed at the journalist.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant abuse terms and tropes featured in abuse directed at the target to help determine the level of risk.
- **Determine if the threat has migrated across platforms** and monitor its spread (including all those platforms where the journalist is present) to help respond to pile-ons which extend the risk of offline harm.
- **Conduct a risk-assessment for the perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened).
- In serious cases, where the threat has multiplied or where a high risk perpetrator is involved, **conduct a network analysis** to determine the original source of the threat, and map its distribution if the capability exists.
- **Indicate if the threat has been reported to the company/platform facilitating the threat and/or law enforcement.** Monitor the progress of these responses.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[63] UN Special Procedures) and monitor follow-up.

63 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

5. Targeted attacks on/or threats against identified family members and close connections (e.g., children)

Research shows that online violence against female journalists **frequently radiates to close family members. In some cases this involves serious threats against children.** When family members are targeted, the risk spreads - along with the pressure and psychological trauma. An international example involves threats against prominent Mexican **Carmen Aristegui**'s family - her young son was targeted and ultimately forced into exile.

Examples from the OSCE region include:

- **Patricia Devlin's** (Northern Ireland) infant child who was threatened with rape via Facebook by a known criminal.[64]
- Serbian journalist **Tatjana Vojtehovski** whose daughter was targeted (both mother and child were subjected to a series of death and rape threats on Twitter).[65]
- US journalist **Kimberley Halkett** (whose teenage daughter was targeted on Instagram).[66]
- Italian journalist **Greta Beccaglia** saw abuse radiate to her family and newsroom during a trial involving a man who assaulted her whilst she was reporting live on TV.[67]
- Spain's **Cristina Fallarás**, whose young children received death threats after she started a Twitter hashtag #Cuéntalo (Tell it) in 2018. In 2021 Cristina left Twitter: "Here I have even endured death threats to my children. ...Twitter is no longer useful to me, it is occupied by males and has become a space for brutal abuse of women with a public presence".[68]

64 op. cit.: Posetti and Shabbir, 2022 (p188).

65 op. cit.: Posetti and Shabbir, 2022 (p60).

66 op. cit.: Posetti and Shabbir, 2022 (p39).

67 Italy: Solidarity with harassed TV reporter Greta Beccaglia, International Federation of Journalists (IJF), 2021: https://www.ifj.org/media-centre/news/detail/category/gender-equality/article/italy-solidarity-with-harassed-tv-reporter-greta-beccaglia

68 Cristina Fallarás: "Es brutal que cueste tanto narrar la violencia machista y siempre haya un macho que te diga 'no puede ser, El Diario, 2019: https://www.eldiario.es/cultura/entrevistas/entrevista-cristina-fallaras_128_1308508.html and Cristina Fallarás también abandona Twitter: "Este lugar acabará siendo un Forocoches", 20Minutos, 2021: https://www.20minutos.es/noticia/4656222/0/cristina-fallaras-abandona-twitter-este-lugar-acabara-siendo-un-forocoches/

- **Arzu Geybullayeva** (Azerbaijan), who experienced threats which radiated to her mother.[69]
- In Spain **María Morán**'s 18-month-old child was targeted as a "bastard child, not recognised by a footballer father", after the sports journalist asked a question in a football press conference in May 2023.[70]
- Dutch journalist **Rena Netjes** went public about online death threats in early 2023 after her family members were targeted.[71]
- **Evgenia Carl** (Slovenia), who brought a lawsuit against then leader of the opposition and future prime minister Janez Jansa for online attacks against her in 2016, said trolls: "attack my children by mentioning them in online articles about me or on social media, exposing them to the public. Nothing, absolutely nothing is sacred to them when it comes to settling accounts with me."[72]

Monitoring guidance for responders

- **Record the threat** (describe the threat, attach a screen grab of the threat, include the URL where relevant, include evidence of any image-based abuse featured e.g., pictures, memes)
- **Identify the secondary target/s of the threat** (i.e., the family member)
- **Identify the medium of the threat** (e.g., social media, chat app, text message, email)
- **Identify the perpetrator** (e.g., username/handle, location [using geolocation information and other forensic techniques], mobile number used, email address, real name and affiliations where evident)
- **Conduct a risk-assessment for the perpetrator** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened)
- In serious cases, where the threat has multiplied or where a high risk perpetrator is involved, **conduct a network analysis** to determine the original source of the threat, and map its distribution if the capability exists.

69 op.cit.: CFWIJ, 2020 and OSCE #SOFJO, 2019.

70 op. cit.: 20 Minutos, 2023.

71 op. cit.: Villamedia, 2023.

72 op. cit.: Global Voices, 2021.

- **Determine if the threat has migrated across platforms.**
- **Monitor targeted abuse across all the platforms** where the journalist and the secondary target/s are present.
- **Identify and monitor any hashtags** being used in connection with the abuse.
- **Indicate if the threat has been reported to the company/platform facilitating the threat and/or law enforcement**, and monitor progress.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[73] UN Special Procedures) and monitor follow-up.

6. *Doxxing as a signal for potential escalation to physical stalking & violence*

Doxxing is the online publication of identifying information associated with a target (e.g., home address, commuting patterns, telephone numbers). It is aided by surveillance technologies but it can also be achieved through manual **stalking/cyberstalking** and presents a very significant risk to the target because of the additional exposure to offline attack that it creates. When a targeted journalist is also doxxed, the act is frequently accompanied by entreaties to digital mobs to 'pile on' and **it can lead to physical stalking and further violence**. Doxxing is a very common feature of online violence campaigns against women journalists and it is a significant risk elevation factor. High profile international cases include **Ghada Oueiss** (Lebanon/Qatar), who was doxxed in a false Facebook account created in her name which included her phone number.[74]

Examples from the OSCE region include:

- **Meri Jordanovska** (North Macedonia) has experienced online abuse since 2009,[75] a TV host called her a 'public enemy' on national TV and he doxxed her on his Facebook page, inviting people to call her for sexual favours.

73 op. cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

74 Posetti et al., ICFJ, 2023: https://www.icfj.org/sites/default/files/2023-02/ICFJ_BigData_Ghada%20Oueiss_Online%20Violence.pdf

75 op.cit.: Birn Balkans YouTube page, 2019.

- **Tanja Milevska**, a freelance journalist from North Macedonia based in Brussels, is threatened online when she writes about geopolitical issues. One perpetrator offered a bounty of 1,000 euros for anyone who could hack into her Twitter account.[76]
- **Daphne Caruana Galizia** (Malta) received threatening phone calls at home. She later found her dogs killed outside her home, and her house was set alight while she and her family were sleeping inside. Ultimately, she was killed in a car bomb blast while driving away from her home.[77]
- **Marianna Spring** (UK) found a message left for her on a board at a train station near her workplace (the BBC in London).[78]
- **Nadine White** (UK) was doxxed on Twitter.[79]
- **Jovana Gligorijević** (Serbia) a journalist at *Vreme*, found her address and personal ID number posted in the comments section of a YouTube video. It was about a 'men's rights activist' being removed from a feminist conference.[80]
- **Sevgil Musaeiva** (Ukraine) was identified in June 2023 - along with her personal information - in a database run by the Ukrainian nationalist website Myrotvorets, and she was falsely accused of using "so-called journalistic activity" to support Russia.[81]
- **Jessikka Aro** (Finland) saw the abusive publication of her private health information in the context of investigating Kremlin-linked trolling networks.[82]
- **Julie Hainaut** (France) was doxxed after publishing an article about a bar in Lyon in 2019.[83]
- **Amy Fenton** (UK) was targeted in May 2020 for her reporting on grooming and crime gangs to such intensity that police estimated there was credible risk to her life and her children's lives and they had to temporarily relocate.

76 op. cit.: CFWIJ, 2020 and Mapping Media Freedom, 2021.

77 op. cit.: Daphne Caruana Galizia Foundation, 2021.

78 op.cit.: Posetti and Shabbir, 2022 (p93).

79 op.cit.: Posetti and Shabbir, 2022 (p154).

80 op.cit.: Posetti and Shabbir, 2022 (p44).

81 op.cit.: CFWIJ, 2023 and CPJ, 2022.

82 op. cit.: NYT, 2016 and BBC, 2017.

83 The Lyon-based journalist was doxxed after writing an article about a colonial-themed bar in the French city - Comment protéger les journalistes contre le harcèlement en ligne, Project Syndicate, 2020: https://www.project-syndicate.org/commentary/french-laws-tackle-online-abuse-of-journalists-by-anya-schiffrin-2020-07/french 2) Council of Europe alert, 2018: https://perma.cc/4RBK-QD7G

- **Hale Gönültaş** (Turkiye) was doxxed in 2018,[84] having reported on Syrian refugees and the humanitarian crisis along the Turkish-Iranian border. In 2022, after Gönültaş documented the horrific abuse of a young woman within a jihadist network in Turkey, her number was posted in a private WhatsApp group. Death threats, phone calls, and a Twitter smear campaign ensued.
- **Brandy Zadrozny** (US) was targeted by a white nationalist, and former speech writer for then-US president Donald Trump, who led an 'open invitation' to dox the NBC reporter in an interview on Fox News' Tucker Carlson show in October 2020; threats were made in "hundreds" of voicemail messages and emails to Zadrozny. She was also told her children were under threat.[85]
- A **New York Times journalist** (US) was doxxed in March 2021 by right-wing news channel One America News Network (OANN). Her phone number was televised in a segment which encouraged viewers to harass the journalist.[86] When OANN later tweeted the segment, they further exposed her number.
- **Cathy Newman** (UK), a Channel 4 news presenter, saw her home address published online after receiving death threats.[87]
- **Katerina Sergatskova** (Ukraine), co-founder of online media Zaborona, reported alleged links between Facebook, a local fact-checking organisation, and neo-Nazi groups. Her private address and photos of her home and five year old son were shared on Facebook in 2020 by a fellow Ukrainian journalist, and she left the country.[88]

84 op. cit. SCF, 2018.

85 op.cit.: Posetti and Shabbir, 2022 (p65).

86 1) op.cit.: Posetti and Shabbir, 2022 (p95) and 2) Twitter Stands By, Lets OANN Link to Reporter's Phone Number, Encourage Users to Harass Her [Updated], Gizmodo, 2021: https://gizmodo.com/twitter-stands-by-lets-oann-link-to-reporters-cell-num-1846509040

87 Cathy Newman interview: Jordan Peterson, a beheading threat and her mosque faux pas, The Times, 2018 (paywalled): https://www.thetimes.co.uk/article/cathy-newman-interview-jordan-peterson-a-beheading-threat-and-her-mosque-faux-pas-wsm5xmwvk and 2) How journalists can hit back at online abuse as C4's Cathy Newman reveals 'totally dehumanising' attacks, Press Gazette, 2021: https://pressgazette.co.uk/news/how-journalists-can-hit-back-at-online-abuse-as-c4s-cathy-newman-reveals-totally-dehumanising-attacks/

88 1) Фейсбук заблокировал Заборону за критику неонацистов. Выяснилось, что украинские фактчекеры соцсети тесно с ними дружат, Zaborana, 2020: https://zaborona.com/ru/stopfake-i-faktcheking-v-facebook/ and 2) CFWIJ Annual Report 2020, The Coalition For Women In Journalism, 2020: https://www.womeninjournalism.org/reports-all/cfwij-annual-report-2020

Monitoring guidance for responders

- **Record the threat/s** (describe the type of threat, attach a screen grab of evidence of the threat).
- **Identify the perpetrator/s** where possible (e.g., username/handle, location [using geolocation information and other forensic techniques], mobile number used, email address, real name and affiliations where evident)
- **Identify the medium/vector/facilitators of the threat** (e.g., social media, chat app, email, text message)
- **Note the identifying information** associated with the doxxing and monitor its spread across channels.
- **Conduct a risk-assessment for the perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened).
- **Monitor targeted abuse across all the platforms** where the journalist is present to help preempt pile-ons connected to doxxing.
- **Identify and monitor any hashtags** being used in connection with the abuse.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers and to measure the speed and spread of attacks and pile-ons, recognising that a doxxing event can precipitate escalation of both online and offline harm.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse any abuse associated with the doxxing episode, the dominant abuse terms and tropes featured in abuse associated with a pile-on directed at a doxxing target.
- **Indicate if the threat has been reported to the company/platform facilitating the threat and/or law enforcement**, and monitor progress.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[89] UN Special Procedures) and monitor follow-up.

89 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

7. Evidence of targeted surveillance &/or interception

The casual deployment of increasingly accessible surveillance software, such as the NSO Group's Pegasus spyware, in online violence attacks against journalists is a significant risk indicator. **Targeted surveillance and interception** can expose female journalists to additional offline harm, including **sexual assault and murder** due to the ability to track and trace the target's movements. It also exposes their sources, family members and colleagues to risk by virtue of the nefarious actor's access to the target's data. Prominent international cases of targeted surveillance (involving Pegasus spyware) include **Carmen Aristegui** (Mexico) and **Ghada Oueiss** (Lebanon).

Examples from the OSCE region include:

- **Khadija Ismayilova** (Azerbaijan) was surveilled through Pegasus spyware for three years (March 2018-May 2021). The senior investigative journalist at the Organized Crime and Corruption Reporting Project (OCCRP) was also targeted by a hidden camera in her bedroom in 2012, with a "sex tape" later published online to shame her, and she spent two years in prison on trumped-up tax charges, before moving abroad.[90]
- **Sevinc Vaqifqizi** (Azerbaijan) had her phone infected with Pegasus spyware in 2019 and 2020 following critical reporting of the Azeri government.[91]
- **Lenaïg Bredoux** (France), a journalist at Mediapart, had her phone infected by Pegasus spyware in 2019 and 2020 after writing about torture in rendition cases in the French and Moroccan governments.[92]
- **Roula Khalaf** (UK), editor of the Financial Times, was targeted by Pegasus spyware in 2018.[93]

90 Khadija Ismayilova profile, Forbidden Stories, 2023: https://forbiddenstories.org/journaliste/khadija-ismayilova/

91 Sevinc Vaqifqizi profile, Forbidden Stories, 2023: https://forbiddenstories.org/journaliste/sevinc-vaqifqizi/

92 Lenaïg Bredoux profile, Forbidden Stories, 2023: https://forbiddenstories.org/journaliste/lenaig-bredoux/

93 Roula Khalaf profile, Forbidden Stories, 2023: https://forbiddenstories.org/journaliste/roula-khalaf/

Monitoring guidance for responders

- **Record the threat/s** (describe the type of threat, attach a screen grab of evidence of the threat e.g., malware, phishing links).
- **Identify the medium of the threat delivery and the type of spyware** if known (e.g., social media, chat app, email, text message), and the potential extent of the surveillance (if possible).
- **Identify the perpetrator/s** where possible (e.g., username/handle, location [using geolocation information and other forensic techniques], mobile number used, email address, real name and affiliations where evident).
- **Conduct a risk-assessment for the suspected perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened).
- **Determine if abuse and harassment towards the journalists experiencing targeted surveillance has migrated across platforms and/or devices** (which might indicate a coordinated attack).
- **Monitor targeted abuse across all the platforms** where the journalist is present to help preempt pile-ons connected to targeted surveillance operations.
- **Identify and monitor any hashtags** being used in connection with abuse
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers and to measure the speed and spread of attacks and pile-ons.
- **Deploy Natural Language Processing techniques**, where a human rights-based capability exists, to analyse the dominant abuse terms and tropes featured in abuse associated with a pile-on directed at a surveillance target.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[94] UN Special Procedures) and monitor follow-up.

94 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018: https://www.osce.org/files/mcdec0003%20safety%20of%20journalists%20en.pdf

8. *Transference of online threats to physical contexts (e.g. physical stalking, being abused in public with disinformation narratives prevalent online; graffiti reflecting online threats)*

Many of the female journalists interviewed by researchers for the UNESCO-ICFJ study, The Chilling, described **being exposed to offline abuse, attacks and harassment that they believed had been seeded online**. And 20% of respondents to the survey conducted for the same study indicated that they **had experienced offline abuse, harassment and attacks that they connected with online violence**. This statistic is an alarming indication of the trajectory of online violence to offline harm, and it underscores the **vital importance of monitoring and recording online violations as a preventive measure**. When online violence spills offline, this is both clear evidence of dangerous risk escalation but also a signal that urgent protective action is required.

Some of the most serious cases in this category involved **physical stalking that began online**. Multiple cases of digital stalking escalating to physical stalking have been recorded by **UK journalists**. A neo-Nazi fake news site published hundreds of disinformation pieces on Finnish journalist **Jessikka Aro**, falsely claiming she had brain damage, spreading conspiracy theories, and calling her a "NATO drug dealer". In this context, she was physically stalked. Pro-Russian activists in Helsinki organized a protest outside the headquarters of her employer, YLE, in 2015, in response to her reporting on a St Petersburg 'troll factory'.[95]

Cristina Fallarás (Spain) experienced offline attacks connected to online violence between October 2019 and June 2021. Her door was graffitied with a cross, and she experienced physical violence on the street such as spitting, being cursed, and pushed. This followed severe abuse on Twitter[96] in the wake of a court case involving five men sentenced for gang-raping an eighteen-year-old girl. Fallarás was called a whore, threatened with rape and death, depicted in a deepfake porn, and logged 1,000 insults a minute online.

At the international level, **Carmen Aristegui** (Mexico) and **Maria Ressa** (the Philippines) saw online threats repeated on flyers or signs outside the newsroom. **Rana Ayyub** (India) had burnt copies of her book dumped at her door during an online hate campaign. **Ana Freitas** (Brazil), who had reported on Gamergate, was targeted with fraudulent deliveries to her home (including packets of worms and gas canisters, food, and sex workers).

Other OSCE region examples include:

- **Greta Beccaglia** (Italy), a sports journalist of local broadcaster Toscana TV, was sexually harassed in 2021, reporting from a football stadium. Footage shows a man assaulting her by violently grabbing her behind, saying that women should not speak about sports, and not about football.[97]
- **Ana Lalić** (Serbia), from Nova.rs, was called a 'mercenary', a 'traitor', and 'unpatriotic'.[98] In this context, she was harassed on the streets, including being thrown out of some establishments on the basis that she was 'not a patriot'.[99]
- **Taylor Lorenz** (US), a frequent target of highly misogynistic online attacks, was punched in the face during a far right rally.[100]
- **Patricia Devlin** (Northern Ireland) received rape and death threats via Facebook Messenger and in graffiti on walls in Belfast.[101]
- **Evgenia Carl** (Slovenia) received envelopes containing death threats and white powder which she said affected her respiratory system. The letters routinely arrive, she said, after court hearings connected to pile-ons caused by the then-leader of the opposition, and future prime minister, Janez Janša.[102]
- **Žydrūnė Jankauskienė** (Lithuania) was abused in a supermarket where she was shopping with her daughter after reporting on corruption.[103] Journalists in South Africa, Sri Lanka, Mexico also reported being yelled at in supermarkets by people spouting abusive terms seeded online.[104]

95 op. cit.: NYT, 2016 and BBC, 2017.

96 op.cit.: El Diario, 2019 and 20Minutos, 2021.

97 op. cit.: IJF, 2021.

98 op.cit.: Posetti and Shabbir, 2022 (p71).

99 op. cit.: Posetti and Shabbir, 2022 (p72).

100 In Charlottesville and elsewhere, U.S. journalists are being assaulted while covering the news, Poynter, 2017: https://www.poynter.org/news-release/2017/in-charlottesville-and-elsewhere-u-s-journalists-are-being-assaulted-while-covering-the-news/

101 op. cit.: Posetti and Shabbir, 2022 (p188).

102 op. cit. Global Voices, 2021.

103 Women journalists and the threats they face: A look across Europe, ECPMF, 2020: https://www.ecpmf.eu/women-journalists-and-the-threats-they-face-a-look-across-europe/

104 1) Women journalists and the threats they face: A look across Europe, European Centre for Press and Media Freedom (ECPMF), n.d.: https://www.ecpmf.eu/women-journalists-and-the-threats-they-face-a-look-across-europe/ and 2) op.cit.: Posetti and Shabbir, 2022 (p93).

Monitoring guidance for responders

- **Record the threat** (describe the type of threat, attach a screen grab/image of the threat if relevant).
- **Identify the location of the threat** which triggered the offline harm (e.g., social media, chat app, email, text message).
- **Identify the location of the associated offline threat/abuse/harassment** and assess the risk according to the level of exposure to harm.
- **Identify the perpetrator** where known (e.g., username/handle, location [using geolocation information and other forensic techniques], mobile number used, email address, real name and affiliations where evident), include any photographic or video evidence of the abuser if available, along with real name and affiliations where evident).
- **Conduct a risk-assessment for the suspected perpetrator** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, or live in close proximity the risk is heightened).
- **Monitor targeted abuse across all the platforms** where the journalist is present to help respond to pile-ons, noting that where online violence escalates to physical context it is also likely to escalate online.
- **Identify and monitor any hashtags** being used in connection with the abuse.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers and to identify linked accounts to monitor as other potential sources of threats linked to offline harm.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[105] UN Special Procedures) and monitor follow-up.

105 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

9. Long-range or large scale attacks with associated risk of significant psychological harm (e.g, networked gaslighting)

Psychological injury caused by online violence is often treated as a lesser harm than other impacts but it can have **devastating consequences, including suicide**, and the types of attack that escalate the risk of psychological injury need to be monitored and appreciated as a trigger for intervention.

Psychological injury (e.g, Post Traumatic Stress Disorder, depression, anxiety etc) **is the most frequently identified impact of online violence** by female journalists internationally according to the ICFJ-UNESCO survey (2020). It manifests physically, with serious implications for the target's health and well-being, including **risk of suicide, their relationships, professional development, and economic security**. The authors' Big Data Case Study on UK journalist Carole Cadwalladr highlighted this threat. In a 2019 interview, she said that being "castigated as a conspiracy theorist and a nutcase", with misogynistic abuse, "...it's very wearing on a day to day basis...our world has now normalised that [abuse]...and you're supposed to put up with it".

Other examples from the OSCE region include:

- **Cathy Newman** (UK), a Channel 4 News presenter who said she felt dehumanised by online death threats and sexualised abuse which were witnessed by her daughter: "I didn't feel like a human being. I felt as if I was being eviscerated by a pack of dogs in the street."[106]
- Polish journalist **Natalia Żaba**, who explained that she found she had trouble "with simple things like paying my bills...I understood that the level of violence I am experiencing every day, whether it's offline, online, doesn't matter. You know, it's pretty [much] the same when it comes to how I feel, and how my body reacts. It's just unacceptable."[107]
- Former *New York Times* journalist **Taylor Lorenz**, who said multiplatform harassment had resulted in: "Weeks where I can't leave my bed and can't function and I'm crying all day, or throwing up all day because of the anxiety and stress that it causes. And there was one point where I definitely didn't want to even live any more."[108]

106 op. cit.: The Times, 2018 and Press Gazette, 2021.

107 op.cit.: Posetti and Shabbir, 2022 (p79).

108 op.cit.: Posetti and Shabbir, 2022 (p79).

- **Evgenija Carl** (Slovenia), who said after being targeted in pile-ons by the future Prime Minister in 2016: "Sometimes I feel depressed and hopeless. Sometimes I feel like I live in a parallel universe because to a normal, reasonable, cultured person, something like this is inconceivable. I wonder how it is possible for 'keyboard warriors' to always be willing to express their thoughts in an aggressive way and how even such a small matter can trigger an explosion of sexism and misogyny."
- **Leona O'Neill** (Northern Ireland) was accused by a blogger of faking the death of Lyra McKee, who was killed in front of her in 2019. She told CPJ: "It impacted my mental health. I became anxious and hyper-vigilant for a time as I dealt with PTSD, not only from the traumatic event I had been a part of, but the tsunami of abuse afterwards. I carried on working and went to trauma counselling, but at times it was extremely difficult to do my job." She has since left journalism.[109]
- **Rianna Croxford** (UK), who said: "I don't think I'll ever really forget that day. The intense anxiety that I felt… I just woke up to hundreds of messages of people criticizing me and abusing me. It did make me question whether I wanted to remain in the profession".[110]

General monitoring guidance:

- **Systematically record threats, abuse and harassment** (describe the type of threat [e.g., rape threat, death threat; threat of other physical violence; threat to harm others; coordinated hate campaign]) against targets at heightened risk of significant psychological injury.
- **Identify the perpetrator/s** (e.g., location [using geolocation information and other forensic techniques], username/handle, include any photographic or video evidence of the abuser if available, along with real name and affiliations where evident)
- **Identify and monitor any hashtags** being used in connection with the abuse.
- **Conduct a risk-assessment for the perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or

109 op. cit.: CPJ, 2019.

110 BBC reporter Rianna Croxford says she nearly quit over abuse 'pile-on' after Kemi Badenoch criticism, The Independent, 2021: https://inews.co.uk/news/media/bbc-reporter-rianna-croxford-nearly-quit-kemi-badenoch-instigated-online-abuse-1076556

if they are associated with a State actor or political leader, the risk is heightened) because the greater the physical risk, the bigger the mental health impacts are likely to be.

- **Identify the threat medium/vector/facilitators** (e.g., social media, chat app, email, text message). Also identify the scale of visibility around the threat (for example, whether more menacing threats are via private Direct Message or are posted to be publicly available and visible on social media platforms). Identify the level of virality of public attacks, which might also be highly relevant as a trigger for more serious psychological impacts.
- **Monitor targeted abuse across all the platforms** where the journalist is present to help respond to pile-ons which exacerbate the risk of psychological injury.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers, and between individual attacks (from one or several perpetrators and over time). Measure the speed and spread of attacks and pile-ons aimed at the journalist, recognising that the mental health impacts of sustained and large scale abuse can be severe.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant abuse terms and tropes featured in targeted harassment to help determine the level of risk to the target's mental health and well-being.
- **Indicate if the threat has been reported to the company/platform facilitating the threat and/or law enforcement.** Monitor the progress of such reports, but also monitor the impact of such processes on the journalist's mental health and well-being, especially if they are required to participate in investigative processes.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[111] UN Special Procedures) and monitor follow-up.

111 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018: https://www.osce.org/files/mcdec0003%20safety%20of%20journalists%20en.pdf

10. The seeding of hashtags and trending narratives associated with judicial harassment, detention & arrest

The seeding of meta-narratives in online violence campaigns against female journalists is linked to legal harassment (particularly in State-linked attacks). The objective is to create an enabling environment through online astroturfing[112] for the persecution, prosecution and conviction of the target. This is clearly evident in the cases of **Maria Ressa**, who refers to this process as "lawfare", and **Rana Ayyub**. In both cases, their names became trending hashtags in association with calls for their arrest e.g., #ArrestMariaRessa, #ArrestRanaAyyub. Both women were also subjected to disinformation narratives suggesting that they were criminals, and corrupt. **Such campaigns are particularly dangerous as they escalate the risk of arrest, detention, prosecution and imprisonment, and mob violence.** In Ressa's case, she was arrested (and later prosecuted and convicted) two years after the #ArrestMariaRessa hashtag first trended on Twitter. Ayyub faces ongoing legal investigations.

Monitoring guidance for responders

- **Record the threat/s** (describe the type of threat, attach a screen grab of evidence of the threat, include evidence of any image based abuse featured e.g., pictures and memes associated with the hashtag)
- **Identify the medium/vector/facilitators of the threat** (e.g., social media, chat app, email, text message).
- **Identify key perpetrator/s** where possible (e.g., username/handle, location [using geolocation information and other forensic techniques], mobile number used, email address, real name and affiliations where evident).
- **Conduct a risk-assessment for the suspected perpetrator/s** who originated the hashtag (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened).

112 The act of manufacturing consent through influence operations designed to create the false impression of a groundswell of support within online communities.

- **Determine if the threat has migrated across platforms.**
- **Monitor targeted abuse across all the platforms** where the journalist is present to help preempt pile-ons connected to 'lawfare' operations.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers and to measure the speed and spread of attacks and pile-ons.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant abuse terms and tropes associated with a hashtag directed at a target.
- **Indicate if judicial harassment follows and links to recorded incidents.**
- **Indicate if the threat has been reported to the company/platform facilitating the threat and/or law enforcement.** Monitor the progress of such reports, but also monitor the impact of such processes on the journalist's mental health and well-being, especially if they are required to participate in investigative processes.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[113] UN Special Procedures) and monitor follow-up.

11. Evidence of coordinated disinformation operations (e.g., repetitive & apparently networked false narratives)

Disinformation purveyors operationalise misogynistic abuse, harassment and threats against women journalists to **undercut public trust in critical journalism and facts in general**. When female journalists are targeted in **disinformation campaigns designed to discredit them professionally or call them into disrepute, false narratives and fraudulent content** (including memes, deep fakes, cheap fakes, spoof accounts etc) proliferate. When the content is spread cross-platform and/or appears to be very similar in style and language, the target is potentially the subject of a **coordinated**

113 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

disinformation campaign. Such campaigns can quickly gain traction and achieve virality, inflicting **significant reputational damage and exposing the target to increased physical and psychological risk**.

41% of female journalists responding to the 2020 ICFJ-UNESCO survey reported that the online violence they experience is associated with coordinated disinformation attacks.[114] Digital disinformation operations are associated with State and foreign State actors, the most serious cases of impunity and legal harassment, and should be understood as a significant indicator for the escalation of online violence.

Examples at the international level include **Rana Ayyub** (India), who has been impersonated on Twitter by perpetrators seeking to expose her to increased risk of physical harm, and misrepresented as a porn star in a deep fake video. While **Ferial Haffajee** (South Africa), **Carmen Aristegui** (Mexico), **Ghada Oueiss** (Lebanon) and **Maria Ressa** (the Philippines) were all subjected to gendered disinformation campaigns.

Examples of this phenomenon in the OSCE region include:

- **Daphne Caruana Galizia** (Malta) was subjected to a sustained and coordinated online disinformation campaign prior to her assassination.[115]
- **Jessikka Aro** (Finland) who investigated pro-Russian Internet trolls in 2014 and uncovered evidence of a State-sanctioned propaganda machine pushing pro-Kremlin narratives through Twitter bots - automated accounts - and bot networks. She became the target of a systematic campaign of cross platform online violence which also moved offline. In a music video campaign against her, she was misrepresented as a hired actress, a "stupid blonde" and NATO spy.[116]
- **Milena Perovic Korac**, from the Montenegrin weekly magazine Monitor has experienced gendered disinformation since April 2011.[117]

114 op.cit.: Posetti et al., 2020.

115 op. cit.: Daphne Caruana Galizia Foundation, 2021.

116 op. cit. NYT, 2016 and BBC, 2017.

117 Female Journalists Attacked Online | Milena Perovic Korac, Birn Balkans YouTube, 2019: https://www.youtube.com/watch?v=U6lk86t30ew

Monitoring guidance for responders

- **Record the threat** (describe the threat, attach a screen grab of the threat, include the URL where relevant, include evidence of any image based abuse featured e.g., pictures, gifs, videos, memes).
- **Identify the medium/vector/facilitators of the threat** (e.g., social media, chat app, text message, email).
- **Identify the perpetrator** (e.g., username/handle, location [using geolocation information and other forensic techniques], mobile number used, email address, real name and affiliations where evident).
- **Conduct a risk-assessment for the perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened).
- In serious cases, where the threat has multiplied or where a high risk perpetrator is involved, **conduct a network analysis** to determine the original source of the threat, and map its distribution via key amplifiers, and to measure the speed of attacks and pile-ons if the capability exists.
- **Identify and monitor the hashtags** being used in connection with the abuse and disinformation narratives.
- **Determine if the threat has migrated across platforms.**
- **Monitor targeted abuse across all the platforms** where the journalist is present.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant abuse terms and tropes featured in the abuse to help determine the main disinformation narratives. Identifying the methods used can also be valuable (typically, coordinated disinformation campaigns leverage platforms' content governance and AI infrastructure). Such analysis could assist with the identification of the key actors by providing insight into their motivation. This data can also aid counter-disinformation work by key responders.
- **Indicate if the threat has been reported to the company/platform facilitating the threat and/or law enforcement.** Monitor the progress of such reports, but also monitor the impact of such processes on the journalist's mental health and well-being, especially if they are required

to participate in investigative processes.

- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[118] UN Special Procedures) and monitor follow-up.

12. *Evidence of orchestrated attacks (e.g., large scale & instantaneous pile-ons)*

When we see **significant spikes in online attacks** in tandem with **high-speed** (sometimes virtually instantaneous) **abuse** via replies to a target's tweets, and/or when network analysis demonstrates connections between abusers, it can be an indication of an **orchestrated attack which can involve central coordination of a network of accounts** (featuring both bots and paid or politically aligned human actors). Such patterns have been detected in the cases of **Maria Ressa** (the Philippines), **Carmen Aristegui** (Mexico) and **Rana Ayyub** (India), for example. They represent a significant risk for escalation to offline harm.

Examples from the OSCE region include:

- **Jelena Obućina** (Serbia), who received sexualised death threats via direct message, said attempts to smear her reputation were part of 'an orchestrated attack by government supporters'.[119]
- **Nastya Stanko** (Ukraine) who reported from the annexation of Crimea for Hromadske in 2014 was targeted in what she and her colleagues described as an 'organized attack' after reporting on the ground in a conflict zone. "A source told us it was the work of three groups of trolls and a bot farm. Though we don't know who commissioned the attack, we do know that their position was strongly pro-government," her colleague Katya Gorchinskaya told the Guardian.[120]

118 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018: https://www.osce.org/files/mcdec0003%20safety%20of%20journalists%20en.pdf

119 1) op.cit.: Safejournalists.net, 2022 and 2) Koalicija za slobodu medija: Procesuirati pretnje upućene Jeleni Obućini, NUNS, 2022a: https://nuns.rs/koalicija-za-slobodu-medija-procesuirati-pretnje-upucene-jeleni-obucini/

120 1) Organization for Security and Co-operation in Europe: The Representative on Freedom of the Media, Dunja Mijatović, 2014: https://www.osce.org/files/f/documents/1/2/127656.pdf and 2) The rise of Kremlin-style trolling in Ukraine must end, The Guardian, 2016: https://www.theguardian.com/world/2016/jul/27/kremlin-style-troll-attacks-are-on-the-rise-in-ukraine-hromadske

- **Jessikka Aro** (Finland). See detailed entry below Indicator 11.
- **Emilia Şercan** (Romania), an investigative journalist who reported on a plagiarism scandal involving a politician, making her the target of harassment and smear campaigns.[121]

Monitoring guidance for responders

- **Record the threat/s** (describe the threat, attach a screen grab of the threat, include the URL where relevant, include evidence of any image based abuse featured e.g., memes).
- **Identify the medium/vector/facilitators of the threat** (e.g., social media, chat app, text message, email).
- **Identify the primary perpetrator/s** (e.g., username/handle, location [using geolocation information and other forensic techniques], mobile number used, email address, real name and affiliations where evident), and other abusers connected to them.
- **Conduct a risk-assessment for the perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened).
- In serious cases, where the threat has multiplied or where a high risk perpetrator is involved, **conduct a network analysis** to determine the original source of the threat, and map its distribution if the capability exists.
- **Identify and monitor the hashtags** being used in connection with the abuse.
- **Determine if the threat has migrated across platforms.**
- **Monitor targeted abuse across all the platforms** where the journalist is present.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers and to measure the speed and spread of attacks and pile-ons.
- **Indicate if the threat has been reported to the company/platform facilitating the threat and/or law enforcement.**

121 1) op. cit.: Reuters, 2022 and 2) Romanian Journalist Emilia Şercan Victim of Smear Campaign, Safety of Journalists Platform COE, 2023: https://fom.coe.int/en/alerte/detail/107637394;globalSearch=true

- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[122] UN Special Procedures) and monitor follow-up.

13. *Misogynistic hate speech (e.g., witch tropes; #presstitutes)*

The use of **misogynistic tropes and sex/gender-based hate speech** is a universal feature of serious online violence campaigns and it is frequently associated with increased risk of offline harm. **Daphne Caruana Galizia** (Malta) was continuously abused as both a "witch" and a "whore" on social media (despite not having active accounts herself, the abuse proliferated on Facebook) prior to her murder. **Carole Cadwalladr** (UK) and **Carmen Aristegui** (Mexico) are also frequently abused as 'witches'.

Another example is the use of the term **'presstitute'** - a portmanteau of 'press' and 'prostitute' - first used in India. It is designed to **discredit female journalists professionally and personally simultaneously**. It also exposes them to increased risk of offline harm in conservative cultures where perceived sexual immorality is punished. We see it (and the hashtag #presstitute) in large abuse detection datasets we have curated in the cases of **Rana Ayyub** (India), **Ghada Oueiss** (Lebanon), **Pauli van Wyk** (South Africa), **Ferial Haffajee** (South Africa), and **Maria Ressa** (the Philippines). Ressa, Ayyub, Oueiss and Haffajee were similarly targeted through misogynistic caricatures. In Ressa's case, her head was depicted attached to a scrotum in viral memes.

Other examples from the OSCE region include:

- **Silvia Bencivelli** (Italy), who was subjected to online violence after a blogger incited dozens of trolls to threaten the freelance journalist in response to a *La Stampa*[123] article where she debunked conspiracy theories. The Perpetrator also posted a YouTube video featuring pictures of the journalist, in which he encourages her rape as payback for being a "misinformer". Bencivelli sued the perpetrator and won.[124]
- Dutch researcher **Rena Netjes**, who 'paid the price' for 'exposing political propaganda' through her research for a documentary on Istanbul terrorist attacks: "I got emails that they made my profile on lesbian/gay/sex accounts".[125]
- **Marianna Spring**, BBC Disinformation Correspondent, is called '"Satan's whore", "the Devil's slut"; she believes her presence onscreen as a broadcast journalist means that "clips that show my face often solicit the most sexist harassment".[126]
- **Jovana Gligorijevic**, a journalist at the liberal Serbian weekly *Vreme*, has been told daily that she is: "a sack of crap that lives in a shop window in the Red Light district," a "vaginal entrepreneur", a "frustrated childless whore" and a "low-paid journalist who occasionally goes to Amsterdam to work as a prostitute to make ends meet".[127]
- **Biljana Blagoeska Petrusheva** (North Macedonia): As a sports journalist, I often face belittling and insults of the type "couldn't you find a man to write about sports", "stay at home and make lunch", "this aunty is so persistent about writing", "you haven't got a clue".[128]

Monitoring guidance for responders

- **Record the threat/s** (describe the threat, attach a screen grab of the threat, include the URL where relevant, include evidence of any image based abuse featured e.g., memes).
- **Identify the medium/s of the threat** (e.g., social media, chat app, text message, email).
- **Identify the perpetrator/s** (e.g., username/handle, mobile number used, email address, real name and affiliations where evident).

122 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018: https://www.osce.org/files/mcdec0003%20safety%20of%20journalists%20en.pdf

123 op. cit:. IJF, 2014.

124 'Così ho portato a processo i complottisti delle scie chimiche che mi minacciavano sul web. E ho vinto', La Repubblica, 2018: https://www.repubblica.it/tecnologia/social-network/2018/04/24/news/processo_scie_chimiche_bencivelli-194705187/

125 op.cit.: Villamedia, 2023.

126 op.cit.: Mapping Media Freedom, 2021.

127 Online abuse now commonplace for Balkan women reporters, BIRN, 2019: https://balkaninsight.com/2019/06/18/online-abuse-now-commonplace-for-balkan-women-reporters/

128 From Normalization to Self-Censorship – Analysis of Online Harassment of Women Journalists in North Macedonia, OSCE, 2022: https://www.osce.org/files/f/documents/b/e/526985.pdf

- **Conduct a risk-assessment for the perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened).
- In serious cases, where the threat has multiplied or where a high risk perpetrator is involved, **conduct a network analysis** to determine the original source of the threat, and map its distribution if the capability exists.
- **Identify and monitor the hashtags** being used in connection with the abuse.
- **Determine if the threat has migrated across platforms.**
- **Monitor targeted abuse across all the platforms** where the journalist is present.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant misogynistic abuse terms and tropes featured in abuse to better understand the scale of the abuse.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers and to measure the speed and spread of attacks and pile-ons.
- **Indicate if the threat has been reported to the company/platform facilitating the threat/s and/or law enforcement**, and monitor progress.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[129] UN Special Procedures) and monitor follow-up.

14. Intersectional abuse (e.g., racism, sectarianism, religious bigotry, homophobia in combination with misogyny)

Intersectional abuse - which occurs at the **nexus of misogyny and other forms of discrimination** - represents a multifaceted risk of increased online violence exposure and impacts, including offline harm. This is a widespread problem

129 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018: https://www.osce.org/files/mcdec0003%20safety%20of%20journalists%20en.pdf

evident in exemplar cases, including: **Ghada Oueiss** (Lebanon/Qatar), a Christian woman practising journalism in conservative Muslim societies; **Rana Ayyub** (India), a Muslim woman being targeted by Hindu Nationalists; **Carmen Aristegui** (Mexico), abused for being a "lesbian" because she "dresses like a man"; **Ferial Haffajee** and **Pauli van Wyk** (South Africa) who are subjected to race-based abuse inflamed by populist political actors.

Other examples from the OSCE region:

- **Alexandra Pascalidou** (Sweden) is targeted for her Greek migrant background.[130]
- Attacks on **Serbian and Macedonian women journalists** include them being called 'Albanian' and 'gypsy' as ethnic insults.
- Black female journalist **Rianna Croxford** (UK), an award-winning BBC Investigations reporter, described a repetitive pattern of racist abuse, including calling her a "monkey" and the suggestion "you've only been hired to fit a quota or tick a box".[131] This is a form of abuse also familiar to The Independent's Race Reporter, Nadine White and VICE UK Editor-in-Chief Zing Tsjeng.[132]
- **Sophie Perry** (UK) launched a network for LGBTQ+ journalists in 2020 but had to leave social media because of the abuse she subsequently sustained online: "People have made references to my sexuality, made very coarse assumptions about my politics, personality and morals… for reporting on publicly available information."[133]
- **Julia Carrie Wong** (US) from Guardian US is targeted on the basis of her Chinese-Jewish heritage and her gender, including through anti-Semitic memes.[134]
- Belgian journalist and novelist **Myriam Leroy** was harassed online for more than nine years by a person who was sentenced to ten months in

130 Alexandra Pascalidou and the neo-Nazi, Deutsche Welle, 2018: https://www.dw.com/en/sweden-alexandra-pascalidou-meets-her-neo-nazi-tormentor/a-42015396

131 op.cit.: Posetti and Shabbir., 2022 (p48) and The Independent, 2021.

132 op.cit.: Posetti and Shabbir., 2022 (p50).

133 Journalist forced to quit social media after suffering homophobic abuse, Hold the Front Page, 2022: https://www.holdthefrontpage.co.uk/2022/news/journalist-forced-to-quit-social-media-after-suffering-homophobic-abuse/

134 op.cit.: Posetti and Shabbir., 2022 (p49).

jail in 2021. The slurs on Twitter were sexist and anti-Semitic.[135]

- **Joanne Chiu** (Canada), a senior reporter at *Toronto Star*, experienced anti-Asian racism during the pandemic, "as a journalist of Chinese descent covering sensitive political issues".[136]
- **Saba Eitizaz** (Canada), a podcast host at *Toronto Star*, received a "horrible message…a filthy, cowardly and pathetic note — the sender didn't leave a real name or contact info and hid behind an encrypted email address".[137]

General monitoring guidance for all responders

- **Record the threat/s** (describe the threat, attach a screen grab of the threat, include the URL where relevant, include evidence of any image-based abuse featured e.g., pictures, memes, videos).
- **Identify the medium/s of the threat** (e.g., social media, chat app, text message, email).
- **Identify the perpetrator/s** (e.g., username/handle, location [using geolocation information and other forensic techniques], mobile number used, email address, real name and affiliations where evident).
- **Indicate the particular intersectional attributes targeted** (e.g., race, religion, sexual orientation).
- **Conduct a risk-assessment for the perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened) but ensure the risk assessment takes on board the heightened intersectional vulnerabilities of the target.
- In serious cases, where the abuse has multiplied or where a high risk perpetrator is involved, **conduct a network analysis** to determine the original source of the threat, and map its distribution if the capability exists.

135 1) Belgium: CFWIJ Welcomes Prison Sentence To Online Abuser Against Journalist Myriam Leroy, CWFIJ, 2021: https://www.womeninjournalism.org/threats-all/cfwij-welcomes-prison-sentence-to-online-abuser-against-journalist-myriam-leroy and 2) Le cyberharceleur de la journaliste Myriam Leroy condamné à 10 mois de prison avec sursis, Rédaction Paris Match Belgique, 2021: https://parismatch.be/actualites/societe/530038/le-cyberharceleur-de-la-journaliste-myriam-leroy-condamne-a-10-mois-de-prison-avec-sursis

136 op. cit.: Toronto Star, 2021.

137 Vicious online attacks won't silence voices in Canadian media, Toronto Star, 2022: https://www.thestar.com/opinion/public_editor/2022/08/11/vicious-online-attacks-wont-silence-voices-in-canadian-media.html

- **Identify and monitor the hashtags** being used in connection with the abuse.
- **Determine if the abuse has migrated across platforms** and track its spread.
- **Monitor targeted abuse across all the platforms** where the journalist is present.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant racist, sectarian, homophobic, transphobic etc terms and tropes featured in abuse to better understand the scale of the abuse and help understand the potential intersectional impacts.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers and to measure the speed and spread of attacks and pile-ons.
- **Indicate if the threat has been reported to the company/platform facilitating the threat/s. and/or law enforcement**, and monitor progress.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[138] UN Special Procedures) and monitor follow-up.

15. State, fake, or partisan media involvement in targeted online violence

The involvement of **State media, State-aligned, heavily partisan, and 'fake news' outlets or journalists/bloggers/influencers** as instigators and amplifiers of online violence abuse can be a feature of coordinated attacks, serving to **escalate, perpetuate and legitimise the cycle of violence**. This type of attack can involve these actors targeting the journalists on social media, or via stories published by their outlets. For example, **Taylor Lorenz** (US), a former *New York Times* reporter, now working for the *Washington Post*, and NBC's **Brandy Zadrozny** (US) were targeted by Tucker Carlson on Fox News in episodes designed to trigger or worsen pile-ons. Similar patterns have been noted by the researchers in the cases of **Carole Cadwalladr** (UK), **Rana Ayyub** (India), **Daily Maverick journalists** (South Africa), and **Maria Ressa** (the Philippines).

138 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

Other examples from the OSCE region include:

- **Jelena Obućina** (Serbia), who was threatened with impalement,[139] was falsely accused by pro-government tabloids Alo and Informer of threatening Serbian president Aleksandar Vučić on television and of inciting anti-state propaganda.[140] She said her words were taken out of context in an attempt to discredit and expose her, increasing the threats she faced online.[141]
- **Arzu Geybullayeva** (Azerbaijan) was the subject of an incendiary opinion article on the site AzLogos, a platform managed by Azerbaijanis living abroad. It claimed that she "hates Azerbaijan and its people", leading to cross-platform abuse on Instagram, Twitter and Facebook, and that she needed to "pay the price" for her "disrespect" and "treason". She was also targeted by a series of tweets from the editor in chief of AzLog.[142]
- **Emilia Şercan** (Romania), an investigative journalist who reported two separate scandals involving senior political figures who plagiarised their doctoral dissertations, was subjected to an orchestrated online smear campaign. The attacks were fuelled by two websites (dezvaluiri.net and oradestiri.net) which posted disparaging articles about Şercan on their respective Facebook pages. In total, 74 sites republished stolen photos taken 20 years earlier by the journalist's fiance and shared articles in relation to them.[143]
- **Žaklina Tatalović** (Serbia), a journalist from independent broadcaster N1, who was targeted with sexually explicit abuse online after the editor-in-chief of a tabloid newspaper took a picture from the journalist's social media account without her consent and published it alongside sexist commentary.[144] A number of Serbian media outlets aligned with the government are accused of initiating disinformation campaigns and judicial harassment against opponents, with women journalists being common targets of misogynistic smear campaigns.

139 op.cit.: Safejournalists.net, 2022.

140 Obućina: Prete mi „predsednikovi ljudi", Danas, 2022: https://www.danas.rs/vesti/drustvo/obucina-prete-mi-predsednikovi-ljudi/

141 The perpetrator was arrested under the Endangering security from Art. 138 st. 3 in connection with para. 1 of the Criminal Code of the Republic of Serbia. ('Tužilaštvo identifikovalo osobu koja je pretila novinarki Jeleni Obućini', NUNS, 2022b: https://nuns.rs/saopstenje-za-javnost-posebnog-tuzilastva-za-visokotehnoloski-kriminal-povodom-pretnji-novinarki-jeleni-obucini/)

142 op.cit.: CFWIJ, 2020 and OSCE #SOFJO, 2019.

143 1) 1) Uncovering a Plagiarism Scandal Among the Romanian Elite, Global Investigative Journalism Network (GIJN), 2022: https://gijn.org/uncovering-a-plagiarism-scandal-among-the-romanian-elite/ and 2) COE, 2023.

144 op.cit.: Posetti and Shabbir, 2022 (p277).

General monitoring guidance for all responders

- **Record the threat/s** (describe the threat, attach a screen grab of the threat, include the URL where relevant, include evidence of any image based abuse featured e.g., pictures, videos, memes).
- **Identify the medium/s of the threat** (e.g., social media, chat app, text message, email, website, TV broadcast).
- **Identify the perpetrator/s** (e.g., username/handle, mobile number used, email address, real name and affiliations where evident).
- **Conduct a risk-assessment for the perpetrator/s** (e.g., if they are identifiable as a person with criminal convictions or connections, or if they are associated with a State actor, the risk is heightened).
- In serious cases, where the threat has multiplied or where a high risk perpetrator is involved, **conduct a network analysis** to determine the original source of the threat, and map its distribution if the capability exists.
- **Identify and monitor the hashtags** being used in connection with the abuse.
- **Determine if the threat has migrated across platforms.**
- **Monitor targeted abuse across all the platforms** where the journalist is present.
- **Deploy Natural Language Processing techniques**, where the capability exists, to analyse the dominant misogynistic abuse terms and tropes featured in abuse to better understand the scale of the abuse.
- **Deploy network analysis and abuse monitoring and visualisation tools** in order to understand connections between abusers and to measure the speed and spread of attacks and pile-ons.
- **Indicate if the threat has been reported to the company/platform facilitating the threat/s and/or law enforcement**, and monitor progress.
- **Indicate if the incident has been reported to an intergovernmental alert system** (e.g., Council of Europe Safety of Journalists Platform, OSCE Ministerial Council Decision No. 3 on the Safety of Journalists,[145] UN Special Procedures) and monitor follow-up.

145 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

SECTION 3:

How to systematically record digital threats

In order to effectively monitor and document online violence against female journalists, it is critical not only to simply record incidents, but to perform both **qualitative and quantitative analysis** of the data and the surrounding context, as described in the following sections of this guidance tool. In turn, this requires the **collection of large datasets, a robust framework for analysis, and technical tools to perform the analysis**. Finally, documentation and monitoring should not be restricted to single data points in time, but should incorporate long-range data in order to understand the progression of abuse over time and better understand causality and escalation of online violence.[146]

While description is important for providing context to an event, adequate monitoring requires achieving the **greatest possible data disaggregation in order to accommodate detail from a wider scope of information and the context within which it is measured**. Simultaneously, it must also meet the requirements for a minimum threshold for systematic and comparable data collection.

A hierarchical classification of incidents allows for compatibility where detailed information is not always available or feasible for all cases of data monitoring across different mediums, countries, data collection possibilities, and attack types.

146 1) The International Protocol on the Documentation and Investigation of Sexual Violence in Conflict highlights the importance of the 'first do no harm' principle for monitoring and recording abuse, 2017: "[P]ractitioners must be fully aware of the possible negative impacts of documentation on victims and other witnesses, the wider community and the investigators themselves; be prepared for the harm those impacts may inflict; and put in place measures to prevent or minimise that harm." and 2) Recognising Sexual and Gender-Based Violence as an Open Source Researcher, Bellingcat, 2023: https://www.bellingcat.com/uncategorized/2023/03/03/sexual-and-gender-based-violence-open-source-researche-osint-digital/?fbclid=IwAR1E5bBuWnZoFD3t6dkfF9_Eb3WT7Fr7WyTek6omcnUTfsaBJRjjAZWtrNE

Monitoring digital threats and attacks entails two primary components:

- **Accurately describing what is taking place** and realistically assessing the scope and nature of the problem.
- Some **degree of evaluation** to make sense of what is going on and to provide context and saliency, for example in the context of goals and targets such as the UN's 2030 Agenda for Sustainable Development.[147]

To this end, **large datasets are required** for the generation of data about the frequency, nature and patterns of abuse being experienced by female journalists, as well as details about the perpetrators and their networks.

Such large datasets cannot be fully analyzed manually, though, because of their sheer size. But automated tools have inherent biases and they are limited in accuracy, since abuse is not always clearly evident, and it can be deliberately subtle, making it tricky to understand even for humans, especially where image-based abuse occurs.

Nevertheless, **human rights-centred AI technology operating with a 'human-in-the-loop' approach** can help to detect and characterize online abuse and ultimately be used to signal potential online violence escalation in real time, thereby enabling rapid response from expert and industry actors (e.g., UN mechanisms, civil society organizations focused on the hybrid safety of journalists, news organizations, and the platforms).

What to monitor and how to record the data

The need for extensive data about online violence against female journalists imposes a number of methodological challenges. In order to strengthen existing analogue monitoring efforts, **a robust and reliable methodology needs to be developed**, encompassing not only the data itself but also the relationships between relevant actors, organizations and events in both time and space. Country and cultural context are also extremely relevant.

This section provides a framework for systematically recording online threats towards female journalists which aligns with international human rights law standards to aid processes of reporting, investigating and prosecuting violations. It is accompanied by a template to help systematize the recording of violations (see Appendix 1), and it includes guidance on:

- What **types of incidents** to monitor
- How to **categorise threats** (using the typology provided)
- Which **platforms and data sources** to target for monitoring
- How to **observe, record and analyze the data** (using both qualitative and quantitative methods).

However, it stops short of providing detailed guidance on the process of big data monitoring and analysis due to the inherent methodological complexity and specialist computational linguistics skills required for such tasks, for which collaboration with experts in Natural Language Processing (NLP) and related AI-powered techniques is recommended.

147 UN's 2030 Agenda for Sustainable Development: https://sdgs.un.org/2030agenda

Categorisation of threats and abuse

Below, we propose a typology which details the types of abuse that should be monitored. **This multi-level classification system allows us to see the abuse at a granular level and within the broader context around incidents of abuse.** Each category is accompanied by a case example drawn from the research-based online violence escalation indicators detailed in Section 2, above.

It should be noted that these categories are not mutually exclusive. For instance, a threat of sexual assault can also be considered as sexually explicit personal abuse, and attacks can involve multiple characteristics such as racism and religious bigotry layered on top of misogyny.

3.1 Abuse typology

Abuse Type	Subtype	Description	Example
Threats		In general, a threat suggests that harm will come to a journalist, or that the journalist deserves harm to come to them	
	Death threat	A threat that suggests the journalist either will or should be killed or die	See detailed 'death threat' case examples under Indicator 1:[148] e.g., Daphne Caruana Galizia (Malta), Jelena Obućina (Serbia), Hale Gönültaş (Turkiye), Arzu Geybullayeva (Azerbaijan), Jovana Gligorijević (Serbia), Tanja Milevska (North Macedonia), Jessikka Aro (Finland), Maria Ressa (the Philippines), Sevgil Musaieva (Ukraine); Tanja Milevska (Belgium/ North Macedonia); Boroka Paraszka (Hungary/ Romania), Sharon Ní Bheoláin (Ireland)
	Sexual assault	A threat that suggests the journalist either will or should be sexually assaulted	See sexual assault case examples under Indicator 1: e.g., Pauli van Wyk (South Africa), Maria Ressa (the Philippines), Rana Ayyub (India), Marianna Spring (UK), Sharon Ní Bheoláin (Ireland), Tímea Karip (Hungary), Ada Borowicz (Poland)
	Physical harm	A threat that suggests the journalist either will or should come to other forms of physical harm/inciting violence	See physical harm case examples under Indicator 1: Tanja Milevska (Belgium/ North Macedonia); Jessikka Aro (Finland)
	Radiation of threats and abuse to close contacts of the target	Online threats which also radiate to hurt and endanger people close to the journalist being targeted, such as children, parents, partners, siblings	See radiation of threats and abuse case examples under Indicator 5: Patricia Devlin (Northern Ireland), Cristina Fallarás (Spain), Kimberly Halkett (US), Rana Ayyub (India), Greta Beccaglia (Italy), Carmen Aristegui (Mexico), Tatjana Vojtehovski (Serbia)

148 The 15 Indicators for online violence escalation are located above.

Abuse Type	Subtype	Description	Example
Attacks on Credibility/ Reputation		Language that implies a person is not fit for their job, cannot be trusted, that their journalism cannot be trusted, or insults towards their intelligence or mental capacity, with the aim of damaging their professional reputation	See Attacks on credibility/ reputation case examples Carole Cadwalladr (UK), Ghada Oueiss (Lebanon), Meri Jordanovska (North Macedonia), Rianna Croxford, Marianna Spring (UK)
Personal Attacks		Personal insults directed towards aspects of the journalist's biological characteristics or identities (sexual, racist, homophobic etc abuse)	
	Misogynistic	Terms that are belittling or degrading to women, or incite hatred towards them	See misogynistic personal attack case examples in Indicator 13: Silvia Bencivelli (Italy); Jovana Gligorijevic (Serbia); Marianna Spring (UK); Rana Ayyub (India); Jessikka Aro (Finland); Rena Netjes (Netherlands); Biljana Blagoeska Petrusheva (North Macedonia); Anthi Pazianou (Greece)
	Sexually explicit	Terms that are sexually explicit or involve sexual acts (may refer to anyone's anatomy)/ photos which are suggestive	See sexually explicit personal attack examples in Indicator 1: Natalia Żaba (Poland)
	Racist	Language that is racist	See racist personal attack examples under Indicator 14: Nadine White, Zing Tsjeng and Rianna Croxford (UK); Serbian female journalists; Alexandra Pascalidou, (Sweden); Julia Carrie Wong, Seung Min Kim (US), Joanne Chiu (Canada)
	Homophobic	Language that is homophobic	See homophobic personal attack case examples under Indicator 14: Sophie Perry (UK)
	General	Other kinds of personal insults (such as use of mild swear words and slurs)	

Abuse Type	Subtype	Description	Example
Belief-based attacks		Attacks based on a person's beliefs or choices	Carole Cadwalladr (UK), Ghada Oueiss (Lebanon), Meri Jordanovska (North Macedonia), Rianna Croxford, Marianna Spring (UK)
	Religious	Language that attacks a person's religion or is derogatory towards their faith (and in some cases, false assumptions about their faith)	See religious belief-based attack case examples under Indicator 14: Myriam Leroy (Belgium), Ghada Oueiss (Lebanon), Rana Ayyub (India)
	Political / sectarian	Language that attacks perceived political affiliations or philosophies	See political / sectarian belief-based attack case examples under Indicator 2: Sevgil Musaieva (Ukraine); Cristina Fallarás (Spain); Jovana Gligorijević, Žaklina Tatalović (Serbia); Boroka Paraszka (Hungary/ Romania)

Table 1: Categorization of abuse types. See Appendix 1 for ideas on implementation when recording incidents of abuse.

As indicated earlier, it is important to situate monitoring within existing frameworks, such as **OSCE Ministerial Council Decision No. 3 on the Safety of Journalists (2018).**[149] At the UN level, Sustainable Development Goal (SDG) indicator 16.10.1 covers "verified cases of killing, kidnapping, enforced disappearance, arbitrary detention and torture of journalists".[150] The addition of "other harmful acts" to categories of violations came via the adoption of a Human Rights Council resolution in 2018.[151] So, we draw on the basic monitoring rationale of SDG 16.10.1 categories of violations against journalists aligned to human rights,[152] which aims to ground the 16.10.1 categories in the measurement of fundamental human rights and to show how 16.10.1 monitoring can be situated within the wider context and practice of human rights. We extend this to include a typology of online violations and their consequences, which are mapped to these categories.

This kind of event-based approach[153] introduces a way of abuse monitoring that not only counts the number of times a specific type of violation has occurred (such as a killing correlated with online violence, or a threat of sexual violence), but also facilitates: the capturing and recording of the severity of the abuse; temporal factors such as long-term and/or repetitive abuse over time and potential escalation to more serious and/or offline threats and attacks; and the effects of the abuse on the targeted individual. **By aligning with existing human rights and SDG frameworks, it provides a monitoring infrastructure that enables States, intergovernmental organizations, civil society groups and news organizations to more systematically capture, record and report violations according to a standardized framework.**

For example, civil society actors conducting shadow monitoring of violations (under SDG 16.10.1) against journalists in national contexts would be able to more easily supplement or present alternative information to official State data reported to intergovernmental bodies like the OSCE, Council of Europe, or UNESCO. **Similarly, a standardized approach to monitoring online violations could also assist news organizations monitoring attacks** on at-risk female journalists to more easily escalate cases, **along with platforms which should be regulated to enforce rapid response within their ecosystems** in the context of serious risk to a target.

149 op.cit.: OSCE Ministerial Council Decision No. 3 on the Safety of Journalists, 2018.

150 United Nations Sustainable Development Goal 16: https://sustainabledevelopment.un.org/sdg16

151 Human Rights Council, The Safety of Journalists. Resolution A/HRC/39/L.7, HRC, 2018: https://undocs.org/en/A/HRC/39/L.7

152 Harrison, J., Maynard, D. and Torsner, S. Strengthening the Monitoring of SDG 16.10.1 and the Manifestations of Violations against Journalists through an Events-based Methodology. Journal of Media and Communication, vol. 8, no. 1: Rethinking the Safety of Journalists, 2020.

153 ibid.

3.2 Online violence incidents mapped to typology of human rights violations

According to the human rights violations typology illustrated below, online violence largely falls under the SDG 16.10.1 category "Other harmful acts". But, critically, the consequences of online violence can be mapped to other violations, including the killing of female journalists (as demonstrated in the cases of Daphne Caruana Galizia in Malta and Gauri Lankesh in India).[154] Until now, the monitoring of "other harmful acts" in this infrastructure has been woefully inadequate,[155] but the potential for online violence to escalate to other forms of harm needs to be properly understood and monitored.

So, below, we have mapped online threats and attacks, and their consequences, to SDG 16.10.1 categories of violations against journalists to human rights violations based on Goetz's conception of human rights derived from international law.[156]

Human Rights Violation	SDG 16.10.1 violation
Violation against the right to life	• Killing
Violation against the right to liberty	• Enforced disappearance • Arbitrary detention • Kidnapping • Other harmful acts

Online violation[157]	Potential consequence of online violation/s
• Threat to murder	• Death • Psychological injury
• Doxxing • Cyberstalking • Targeted surveillance • Astroturfing campaigns designed to legitimize 'lawfare' against a target	• Kidnapping • Detention • Physical or sexual attack • 'Lawfare' • Psychological injury

154 In the age of false news: a journalist, a murder, and the pursuit of an unfinished investigation in India, Forbidden Stories, 2023: https://forbiddenstories.org/story-killers/gauri-lankesh-in-the-age-of-false-news/

155 op.cit.: Harrison et al., 2020.

156 Goertz, G. Social Science Concepts: A User's Guide. Princeton University Press, 2006.

157 These violations should be registered in accordance with human rights standards which allow only very narrow exceptions in limited contexts.

Human Rights Violation	*SDG 16.10.1 violation*
Violation against personal dignity	• Torture • Other harmful acts
Violation against the right to privacy	• Other harmful acts

Online violation	*Potential consequence of online violation/s*
• Doxxing • Cyberstalking • Surveillance • Threats of sexual violence • Threats of physical violence • Abuse that is sexually explicit, racist, homophobic, sectarian, bigoted	• Physical stalking • Sexual violence • Physical violence • Psychological injury
• Doxxing • Cyberstalking • Surveillance • Hacking • Unwanted private messages • Non-consensual sharing of intimate images/video	• Physical stalking • Increased risk of physical and sexual assault • Increased risk of murder • Psychological injury

Human Rights Violation	*SDG 16.10.1 violation*
Violation against the right to expression	• Other harmful acts

158 "In the Philippines, this is the practice of publicly labelling individuals and organizations as "enemies of the state", "communist terrorists", or "members of communist front organizations". "Red-tagged" journalists become open targets for violence online and offline. Recent killings of human rights activists have also been linked to this practice, highlighting the risks facing women journalists who are targeted in this way.. 'Red-tagging' involves falsely identifying a journalist as a member of the Communist Party of the Philippines (CPP) or the party's military wing, the New People's Army (NPA). Individual community journalists are more vulnerable to these attacks but in some cases entire news outlets have been tagged as being the media of the CPP-NPA'. op.cit.: Posetti and Shabbir, 2022 (p294).

Online violation	*Potential consequence of online violation/s*
• Threat to murder • Threats of sexual or physical violence • Misogynistic abuse and threats • Intersectional abuse • Hate speech • Doxxing • Cyberstalking • Harassment • Sexual harassment • Surveillance • Non-consensual sharing of intimate images/video • Defamation • Insult • Slander • Libel • Hate speech • Professional and reputational threats	• Physical stalking • Increased risk of physical and sexual assault • Increased risk of murder • Psychological injury • Censorship • Self-censorship • Publication restrictions • Prohibition on speech and deplatforming • Red-tagging[158]

Human Rights Violation	*SDG 16.10.1 violation*
Violation against the right to association and assembly	• Other harmful acts
Violation against the right to movement and residence	• Other harmful acts
Violations against the right to own/retain property	• Other harmful acts
Violations against the right to protection of reputation	• Other harmful acts
Violations against the right to freedom from discrimination	

Online violation	*Potential consequence of online violation/s*
• Misogynistic abuse and threats • Intersectional abuse • Abuse connected to the demonization of journalism	• Self-censorship which leads to withdrawal from online communities • Deplatforming of journalists targeted in fraudulent mass-reporting of accounts from online communities
• Doxxing • Surveillance	• Closure of an office • Restricted movement • Forced relocation • Exile
• Defamation • Insult • Slander • Libel • Professional and reputational threats	• Financial hardship (loss of employment) • Reputational damage affecting professional standing
• Hate speech • Sexual discrimination • Gender-based discrimination • Racial discrimination • Religious discrimination • Political discrimination • Other forms of discrimination	• Incitement to hatred • Hate crimes • Mob violence

Human Rights Violation	*SDG 16.10.1 violation*
Violations against the right to integrity	• Other harmful acts

Online violation	*Potential consequence of online violation/s*
• Death threats • Threats of physical violence • Threats of sexual violence • Financial threats • Threats against family members • Intimidation • Harassment • Sexual harassment • Use of abusive or hateful language • Gendered vitriol • Image and video-based abuse • Deep fakes • Shallow fakes	• Physical assault • Sexual assault • Murder • Psychological injury • Economic hardship • Radiating harm to family members and colleagues

Table 2: Mapping of online threats and attacks, and their consequences, to SDG 16.10.1 categories of violations against journalists to human rights violations based on Goetz's conception of human rights derived from international law.[159]

This proposed scheme aims to contribute to **building a more solid evidence-base** from which it is possible to understand the background of intensified attacks against journalists and to support the development of effective measures to mitigate and prevent these. It builds on recent recommendations for monitoring violations against journalists more widely, in line with SDG 16.10.1, which focus primarily - though not exclusively - on offline violations.[160] Those recommendations include the development of an event-based approach by means of methods and tools that aim to strengthen ongoing monitoring efforts on a wide range of violation types, along two primary paths.

159 op.cit.: Goetz, 2006. See Appendix 1 for ideas on implementation when recording incidents of abuse.

160 op.cit.: Harrison. et al., 2020.

First, it is critical to improve existing monitoring through the use of better and more consistent data. This applies equally, in the case of online violence, to newsrooms employing journalists at risk. Second, **data analysis tools, including Natural Language Processing should be used**, where available, to extract relevant contextual information and facilitate a more comprehensive analysis of the data.

Furthermore, the construction of a solid evidence base can be strengthened by **considering the legal implications of individual violations**. Our recommendations include clearly linking the different categories of violations to international law and methodological standard and monitoring practices, such as those developed by the Office of the High Commissioner for Human Rights (OHCHR) and other international mechanisms, including the International Classification of Crime for Statistical Purposes

(ICCS) disseminated by the UN Office of Drugs and Crime (UNODC).[161] This is in line with the UN's 2030 Sustainable Development agenda which not only includes legal definitions of the aspects of the key violation categories of killing, kidnapping, enforced disappearance, arbitrary detention, and torture, but also correlates them with the ICCS criminal codes.[162]

For example, the category of 'killing' is defined as: *"any extrajudicial execution or other unlawful killing by State actors or other actors acting with the State's permission, support or acquiescence that were motivated by the victim, or someone associated with the victim, engaging in activities as a journalist, trade unionist or human rights defender; or while the victim was engaged in such activities; or by persons or groups not acting with the support or acquiescence of the State whose harmful acts were either motivated by the victim engaging in activities as a journalist, trade unionist or human rights defender, and/or met by a failure of due diligence on the part of the State in responding to these harmful acts, such a failure motivated by the victim or associate engaging in activities as a journalist, trade unionist or human rights defender; and other unlawful attacks and destruction in violation of international humanitarian law leading to or intending to cause the victim's death, corresponding to ICCS codes 0101, 0102 and 110139..."*[163]

We note, however, that mappings between categorisation schemes are not straightforward: the ICCS scheme includes **the concept of intentionality**, so that for example one subcategory of 'killing' is defined as: "intentional homicide related to political agendas, including killings by terrorist groups with a political agenda, political assassination, and targeted killing of journalists for political reasons". Similarly, not all killings are classified under the top-level ICCS code of 01 (which is defined as "Acts leading to death or intending to cause death") since killings related to "war crime" are classified under the code of 11 (defined as "Other criminal acts not elsewhere defined") which itself has a subcategory "Unlawfully killing, causing or intending to cause death or serious injury associated with armed conflict". **These classifications and definitions are important when it comes to legal redress for online violence and its escalation.**

161 International Classification of Crime for Statistical Purposes (ICCS), 2015: https://www.unodc.org/documents/data-and-analysis/statistics/crime/ICCS/ICCS_English_2016_web.pdf

162 Berger, G, 2020. "New opportunities in monitoring safety of journalists through the UN's 2030 sustainable development agenda." Media and Communication 8.1: 78-88: https://d-nb.info/1226487696/34.

163 UN Stats, SDG indicator metadata, 2018: https://unstats.un.org/sdgs/metadata/files/Metadata-16-10-01.pdf

3.3 Online violence and the International Classification of Crime for Statistical Purposes

The table below shows a mapping of the most relevant ICCS codes to the categories of online violence and their consequences as developed above. **Note that not all violations of human rights, and not all types of online violence, can be categorized as crimes, so not every violation will be associated with a ICCS code.** Since the ICCS codes have a complex system of subcategorization, here we select the most relevant code.

According to the ICCS classification, **threats to commit a crime can also be considered relevant**. This is an important factor for monitoring online violence towards women journalists which routinely involves threats of violence. Following the ICCS guidelines, threats to killing, assault etc. are therefore attached to the same ICCS codes by our mapping as the acts themselves, with the exception of "threat of physical assault" which has its own subcategory of "assaults and threats" in the classification.

For any available dataset, further data descriptors should be made available to facilitate the interpretation of statistical data. While most of the crimes, and their statistical reporting, refer to offences actually committed by one or more direct perpetrators (whether known or not), data can also include cases of threats to commit a certain crime or when the offence consisted of planning or assisting others to commit it. It is therefore important that information be provided about whether available data on criminal offences (and perpetrators) include or exclude the following behaviours in the counts for the categories:

- threats to commit the crime
- aiding/abetting/accessory to the crime
- accomplice to the crime
- conspiracy/planning the crime
- incitement to commit the crime.

This information should ideally be captured and stored for every criminal offence to indicate whether the recorded event refers to a threat, a case of aiding/abetting/accessory to the crime or any other typology in the list above. In such cases, the desired statistical outputs can be produced by either including or excluding such events from the aggregate counts. Alternatively, the information on the inclusion of such cases can be provided at an aggregated level of crime categories, in the form of meta-data.

Violation/ consequence	ICCS code	Description of ICCS code
Killing	01	Acts leading to death or intending to cause death
Physical attack/ violence	0201	Assaults and threats
Threats of attack/violence	02012	Threat
Kidnapping	020221	Kidnapping
Detention	02022	Deprivation of liberty
Psychological injury	0208	Acts intended to induce fear or emotional distress
Harassment	02081	Harassment
Physical stalking/ cyberstalking	02082	Stalking
Defamation	0209	Defamation or insult
Insult	0209	Defamation or insult
Slander	0209	Defamation or insult
Libel	0209	Defamation or insult

Violation/ consequence	ICCS code	Description of ICCS code
Unwanted private messages	02011	Invasion of privacy
Non-consensual sharing of intimate images/ video	0201	Assaults and threats
Doxxing	02111	Invasion of privacy
Surveillance	02111	Invasion of privacy
Sexual attack/ violence	0301	Sexual violence
Sexual harassment	030122	non-physical sexual assault
Astroturfing	0709	Other acts involving fraud, deception or corruption
Censorship / self-censorship	0803	Acts related to freedom of expression or control of expression
Publication restrictions	0803	Acts related to freedom of expression or control of expression
Prohibition on speech and deplatforming	0803	Acts related to freedom of expression or control of expression

Violation/ consequence	ICCS code	Description of ICCS code
Red-tagging	01	Acts leading to death or intending to cause death
Deplatforming of journalists	0803	Acts related to freedom of expression or control of expression
Hate speech	080322	Violations of norms on intolerance and incitement to hatred

Table 3: Mapping between violations and International Crime Classification Codes.

In line with these recommendations, building an evidence base for monitoring online violence against female journalists thus requires a **strong theoretical grounding in a comprehensive typology of violations, and digital tools to identify and analyze both the attacks and threats themselves** as they occur on social media, but also the temporal and contextual information necessary for understanding the pathways of online violence escalation.

Platforms and data sources

While it is important to remove the onus from female journalists to manage responses to the online violence they experience, recognizing that **it is not sustainable for them to continue being both the** primary targets and first responders,[164] in incidents of online violence against them must be recorded and sometimes the journalist themself is the only person in the position to do so 'in the moment' (especially where these threats are issued via direct messages and chat apps). They can also be appropriate bystander recorders of attacks on other journalists. Online violence experienced on any digital source - including social media, chat apps, mobile phone, online fora, email etc - and information about the medium, the platform, and methods of attack should be recorded wherever possible. Where data is publicly accessible, such as on

164 op.cit. Posetti. et al., 2021.

Twitter, large scale data can be collected e.g., all messages to and from a journalist, so that the abuse can be analysed in context and over time.[165] **However, technical expertise, privacy issues around gathering and storing data, and the terms and services of most social media platforms prevent this being feasible in many cases.**

General guidelines for monitoring the data revolve around **recording relevant material rapidly and clearly**, so that the nature and context of the threat is clear to others observing and analyzing the data at a later date. Both descriptive and statistical information are critical, so **data should be recorded and analyzed both qualitatively and quantitatively**. The exact nature of the recording and analysis may vary depending on the type of threat (see Section 5 for detailed guidance) but we provide some general principles below. In both cases, once the information has been recorded, where possible, Natural Language Processing (NLP), network analysis and visualization tools should be deployed to analyse the data in more detail. This includes **looking at dominant abuse terms and topics** featuring in attacks, **investigating connections** between abusers, and **measuring the speed and spread of attacks**, along with the methods (e.g., use of threats embedded in image-based abuse).

Qualitative data

In the first instance, threats and abusive material should be recorded as potential evidence as soon as possible. This could be done by a trusted competent person where the mental health impacts mitigate against the journalist themself taking responsibility for this task. **A screenshot of the message should be captured in case it is later deleted by the author, or the platform.** Message threads should also be recorded where evidence of pile-ons or support for the threat is found. Metadata about the threat should also be recorded, including timestamp and medium (e.g. social media platform, and whether the message is private, within a closed group, or public) and all available information about the sender (e.g., username/handle, mobile number used, email address, real name and affiliations where evident).

The threat should be described in order to explain the context, since this may not be evident to others (for example, a manipulated image, synthetic

165 At the time of publication, Twitter had just been rebranded as 'X', and API access to massive Twitter datasets had become prohibitively expensive. Nevertheless, academics, civil society organisations and news outlets were exploring alternative methods of big data collection from the platform.

media like deep fakes, oblique references to prior or upcoming events or messages, any relevant contextual situation) and any URLs to or screenshots of the message should be provided.

Quantitative data

In order to understand the scale and severity of attacks, quantitative data is required. **This should include not just absolute numbers but also relative numbers** (for example, if all messages sent to a journalist can be collected, metrics can be calculated such as percentages of messages which are abusive).

Useful statistics to include are:

- Number of abusive messages, broken down by type of abuse
- Percentage of all replies and messages to a journalist that are abusive (normalizes the abuse by volume of messages)
- Timeline distribution of abusive tweets (indicative of abuse spikes related to specific events)
- Statistics relating to timing of abusive messages sent in reply to a tweet (e.g. abuse frequently sent within a few seconds of a message being posted may be indicative of bots or networked abuse)
- Topics connected with the abuse sent (may indicate useful context around the abuse)
- Details of most frequent abuse senders

Statistics such as these, especially in conjunction with visualisations such as charts and graphs, **can help understand the bigger picture** and identify instances of orchestrated attacks or particularly threatening or worrying behavior which has the potential to escalate. For example, when we see significant spikes in online attacks in tandem with high speed (sometimes virtually instantaneous) abuse via replies to a target's tweets, and network analysis demonstrates connections between abusers, it can be an indication of an orchestrated attack.

Analysis of abuse must always be viewed in context - for example, the political situation, the topics covered by the journalist and the kinds of messages that provoke abuse all give insight into the situation. Relevant timelines of events which correlate with abuse spikes are a good example of understanding the bigger picture.

APPENDIX 1:

A template for recording violations

A model template for recording digital violations against female journalists (which could also be adapted for other high risk targets) and instructions for its use is included below. It includes mock entries to help guide implementation.

The template below provides instructions for using the monitoring template (page 97) which supports implementation of the OSCE RFoM Online Violence Monitoring Guidelines. The template also includes mock entries as guidance.

Note: As much information as possible should be recorded about a violation so that a record is made of the abuse - providing evidence of what happened and its impacts can help track and consolidate abuse and details of its perpetrator.

Note: Always save a copy of any messages (including URLs) and/or take screenshots as soon as possible, in case messages are later deleted. It's wise to keep a folder to store copies for later reference.

Label	Details to provide	Explanation
Date and time of violation	Date and time (refer to time stamp associated with the Online Violation)	This adds vital context and aids the process of monitoring over time
Identity of target	Name, handle, designation, contact details	This allows systematic monitoring of individuals being targeted and aids risk management
Online violation type	See page 100 for examples of online violation types to record in the monitoring template. This lists the types of violations mapped to international human rigths standards, according to Table 2 (see page 70, section 3.2) of the OSCE RFoM Monitoring Guidelines	Identifying the type of abuse in this way helps monitors and responders to understand the nature and severity of the abuse. It can also assist with the process of escalating complaints and threat alerts. It is possible that an incident will involve multiple online violation types. In that case, select the most serious violation from the menu and add relevant additional details to the 'context' or 'notes' columns of the template
Abuse type	See abuse typology provided in Table 1 (on page 64, in section 3.1) of the OSCE RFoM Monitoring Guidelines	Labeling the type of abuse (e.g., sexism/misogyny, racism, homophobia/transphobia, antisemitism, religious bigotry, intersectional abuse) helps to provide essential context to understand the nature of the abuse and its potential impacts
Perpetrator identity	Name, handle, location, mobile number, email, affiliations (professional/institutional)	This is important for tracing and monitoring attacks across a group of targets and escalating cases with law enforcement. This will help to action a response, especially if there's a threat of physical safety associated with it, or if the perpetrator is a person with known criminal convictions, or is known to the individual
Medium (social media platform/ email/chat app/ text message)	Identify the social media platform, email provider, chat app service or text message carrier	It is essential to identify the site of the violation to trace the movement of abuse across platforms and to seek redress
Evidence of violation	Screenshots, images, URLs	This will aid reporting of violations and the escalation of complaint while also helping to track and trace similar threats
Additional data associated with the violation	Hashtags and/or other handles, email addresses or phone numbers connected to the incident	Associated information such as hashtags or memes that frequently occur in connection with the abuse are all useful for analysing the bigger picture

Label	Details to provide	Explanation
Online Violence Escalation Indicators (1-15)	Code the violation from the dropdown menu according to the descriptions of the 15 online violence escalation indicators from the OSCE RFoM Online Violence Monitoring Guidelines	This will aid risk assessment. Where an attack triggers multiple indicators select the most urgent from the menu in the template (tab #2) and add relevant additional details to the 'context' or 'notes' columns of the template. When an indicator is triggered, follow the instructions in the OSCE RFoM Monitoring Guidance for monitoring applicable to specific indicators
Context	Indicate, for example, if this violation was part of a broader attack; specific trigger; presence of intersectional abuse)	Details about the context in which the abuse occurred, e.g. in response to a particular article, as part of a wider attack, and relevant information to help analysts/responders better understand the abuse and its spread are valuable
Risk assessment	How serious is the risk of offline harm: High/ Medium/Low	This will help appropriately calibrate responses to the urgency of the risk and allow others to understand why action needs to be taken, and how urgent that action might be
Describe the impact/s of the incident/s on the individual	e.g., Psychological impacts, economic impacts, impacts on other family members, colleagues or sources	This will allow responders to focus on the needs of the individual and monitor others to whom the abuse might radiate
Threat reported to law enforcement	YES/NO, please identify police service/authority + date of report + contact details	This will help track cases that have been reported and escalated with law enforcement
Threat reported to the platform/s	YES/NO, please identify the platform + date of report	This will help track cases that have been reported and escalated with the platform
Threat reported to relevant journalism safety/press freedom organization/s	YES/NO, please identify the entity + date of report + contact details	This will help track cases that have been referred to advocacy organisations

Label	*Details to provide*	*Explanation*
Threat reported to intergovern-mental organisation	YES/NO, please identify the entity + date of report + contact details	This will help track cases that have been referred to IGOs, potentially aid coordination, and help inform IGOs about the online violence-offline harm nexus
Supervisor of target	Name, title, contact details	This will aid organizational efforts to monitor the case
Protective action/s taken by key responders	Including provision of physical/digital security, psychosocial support etc	This will help the organization remain focused on responding to the needs of the target
Date of follow-up	Indicate when follow-up action (e.g., escalating the case with law enforcement, the platforms etc) occurred	This will assist with monitoring progress of the case over time and encouraging accountability on the part of key responders. This will help with resolution if no action has been taken by the relevant bodies
SDG 16.10.1 Violation type	Code the violation from the dropdown menu featuring the relevant Sustainable Development Goal indicators described in the OSCE RFoM Online Violence Monitoring Guidelines	This will help with monitoring and reporting at the IGO level and aid the systemization of recording violations
ICCS Code	The violation will be automatically categorised according to the relevant International Classification of Crime for Statistical Purposes (ICCS) described in the OSCE RFoM Online Violence Monitoring Guidelines	This will help with monitoring and reporting at the IGO level and aid the systemization of recording violations while also assisting criminal investigations associated with online violence
Additional notes	Record any additional observations (e.g., indicate when multiple indicators have been triggered)	This will support investigations and allow for the observation of correlations

Template with examples

Date and time of violation	18.04.23	24.04.23
Identity of target	Tamara Neugerbauer	Jane Smith
*Online Violation Type**	Death threat	Non-consensual sharing of intimate images / video
Abuse type	Select an option	homophobia / transphobia
Perpetrator identity	@IncelWarrior	Anonymous
Medium (social media platform/ email/chat app/text message)	Facebook	Facebook Messenger, porn sites
Evidence of violation	Screenshot from Facebook	Screenshot from Facebook Messenger
Additional data: Hashtags and/or other handles, email addresses or phone numbers associated with abuse	#liaridiot23	@guapismenti55009
*Online Violence Escalation Indicators (1-15)***	1 Death/rape threats	6 Doxxing
Context	The day after she published an investigation about a leading politician's alleged links to corruption	3 of the target's private pictures were hacked and leaked
Risk assessment (How serious is the risk of offline harm: High/Medium/Low)	High	High

Describe the impact of the incident/s on the individual	Terrified. Had to take time off work	Shock, feeling of being belittled, cowed, violated
Threat reported to law enforcement	YES / national police service / 18.04.22 / police@template.com	YES / national police service / 24.04.22 / police@template.com
Threat reported to the platform/s	YES / Twitter/18.04.22	YES / Facebook / 18.04.22
Threat reported to relevant journalism safety/press freedom organization	YES / national press freedom organisations / 19.04.22 / person@digitalsecurityorganisation.com	YES / national press freedom organisation / 19.04.22 / person@digitalsecurityorganisation.com
Threat reported to intergovernmental organisation	NO	NO
Supervisor of target[166] *(Name, title, contact details)*	Jonathan Schmidt	Samantha Taylor
Protective action/s taken by key responders	Relocated for protection Psychological support provided Digital security checks	Digital security consultant reviewed the target's exposure and privacy settings
Date of follow-up	20.04.22	30.04.22
SDG 16.10.1 Violation type[167]	Other harmful acts	Other harmful acts

166 Where one exists. For freelancers, we recommend contacting experts at civil society organizations.

167 6.10.1 Categories: (note that most of these don't explicitly encompass online abuse, but they do include threats and 'other harmful acts' which can capture targeted online violence)
- Killing
- Kidnapping
- Enforced disappearance
- Arbitrary detention
- Torture
- Other harmful acts

ICCS Code[168]	
Additional notes	
Threats of attack / violence: 02012	Harassment: 02081
We are continuing to check that we have engaged the services of the most suitable workplace psychologist for this journalist	Considering resources and providing journalist with a new work-only phone device, separate to private one

*Online violation types used in column "Online Violation Type"

- Death threat
- Threat of sexual attack
- Threat of other physical attack
- Sexual harassment
- Cyberstalking
- Defamation/slander/libel
- Unwanted private messages
- Non-consensual sharing of intimate images/video
- Doxxing
- Surveillance
- Astroturfing
- Hate speech

**List of indicators used in column "Online Violence Escalation Indicators"

1. Death/rape threats
2. State/foreign State actor/political extremist involvement
3. Proximity to attackers
4. Threats associated with impunity cases
5. Attacks on family members etc.
6. Doxxing
7. Surveillance/interception
8. Transference to physical contexts
9. Long-range / large scale attacks
10. Hashtags/narratives relating to detention, arrest, etc.
11. Evidence of coordinated disinformation
12. Evidence of orchestrated attacks
13. Misogynistic hate speech
14. Intersectional abuse
15. State, fake, or partisan media involvement

168 See page 82 for ICCS codes.

APPENDIX 2:

Resources and organizations providing assistance

Here, we present a non-exhaustive curation of applied research, resources and services designed to support women journalists experiencing online violence and aid efforts to improve responses to the crisis.

- The Chilling: A global study of online violence against women journalists (2022)

  Editors: Julie Posetti and Nabeelah Shabbir (ICFJ & UNESCO)

  Drawing on 15 country case studies, 182 interviews, 700+ survey respondents and analysis of 2.5 million social media posts, this groundbreaking study includes a thematic analysis of 10 global trends in gender-based online violence, a taxonomy of 12 globally recognizable types and methods of attack to prepare for, actor specific assessments, and 106 recommendations for action in response to 35 key findings that point to the need for responses to online violence to be strengthened in technological sophistication and collaborative coordination. It also features a 25-step tool for developing online violence responses that respect freedom of expression.

  https://www.icfj.org/our-work/chilling-global-study-online-violence-against-women-journalists

- The Chilling: Assessing Big Tech's Response to Online Violence Against Women Journalists (2022)

  Authors: Julie Posetti, Kalina Bontcheva and Nabeelah Shabbir (UNESCO)

  https://unesdoc.unesco.org/ark:/48223/pf0000383044

- The Chilling: What more can newsrooms do to combat gendered online violence? (2022)

  Authors: Julie Posetti and Nabeelah Shabbir (UNESCO)

  https://unesdoc.unesco.org/ark:/48223/pf0000383043.locale=en

- The Chilling: Legal and normative frameworks for combatting online violence against women journalists (2022)

  Authors: Angelique Lu, Julie Posetti and Nabeelah Shabbir (UNESCO)

  https://unesdoc.unesco.org/ark:/48223/pf0000383789

- The Chilling: Global Trends in Online Violence Against Women Journalists (2021)

  Authors: Julie Posetti, Nabeelah Shabbir, Diana Maynard and Kalina Bontcheva (UNESCO)

  A discussion paper previewing the full study published in 2022.

  https://www.icfj.org/sites/default/files/2021-04/The%20Chilling_POSETTI%20ET%20AL_FINAL.pdf

- Online Violence Against Women Journalists: A Global Snapshot of Incidents and Impacts (2020)

  Authors: Julie Posetti, Nermine Aboulez, Kalina Bontcheva, Jackie Harrison & Silvio Waisbord (UNESCO)

  Findings from a global ICFJ-UNESCO survey of over 700 women journalists.

  https://www.icfj.org/sites/default/files/2020-12/UNESCO%20Online%20Violence%20Against%20Women%20Journalists%20-%20A%20Global%20Snapshot%20Dec9pm.pdf

- Maria Ressa: Fighting an Onslaught of Online Violence (2021)

  Authors: Julie Posetti, Diana Maynard and Kalina Bontcheva (ICFJ)

  A groundbreaking big data case study examining over half a million social media posts directed at the Nobel Laureate.

  https://www.icfj.org/our-work/maria-ressa-big-data-analysis

- Rana Ayyub: Targeted online violence at the intersection of misogyny and Islamophobia (2023)

  Authors: Julie Posetti, Kalina Bontcheva, Hanan Zaffar, Nabeelah Shabbir, Diana Maynard, and Mugdha Pandya (ICFJ)

  A big data case study on the award-winning Indian journalist.

  https://www.icfj.org/sites/default/files/2023-04/Rana%20Ayyub_ICFJ_Case%20Study.pdf

- Ghada Oueiss: A journalist at the epicenter of online risk amid weaponized geopolitical threats (2023)

  Authors: Julie Posetti, Diana Maynard, Aida al-Kaisy, Zahera Harb and Nabeelah Shabbir (ICFJ)

  A big data case study on the Al Jazeera Arabic principal presenter.

  https://www.icfj.org/sites/default/files/2023-04/Ghada_ICFJ_Case%20Study.pdf

- #SOFJO Resource Guide, "Walk the talk: What key actors can do for the safety of female journalists online" (2020)

  Authors: Dr. Silvia Chocarro, Sarah Clarke, Paulina Gutiérrez and Judy Taing, OSCE Representative on Freedom of the Media published in Русский, Albanian, Қазақша, Кыргызча, Македонски, Serbian, Tajik, Turkish, Uzbek.

  https://www.osce.org/representative-on-freedom-of-media/471903

- A Dark Place (2018)

  Director: Javier Luque, OSCE Representative on Freedom of the Media and the International Press Institute (IPI)

  "First-hand experiences shared by leading women journalists targeted with online violence."

  https://ipi.media/documentary-film-a-dark-place/

- A Perfect Propaganda Machine (Hungary Report), (2023)

  By Lucina Di Meco and Sarah Hesterman, #ShePersisted

  Part of five country-specific case studies in a #MonetizingMisogyny research series analysing the - patterns, impacts and modus operandi of online attacks and disinformation campaigns targeting women leaders.

  https://she-persisted.org/our-work/research-and-thought-leadership/

- A Digital ResilienceToolkit for Women In Politics: Persisting and Fighting Back Against Misogyny and Digital Platforms' Failures (2022)

  Author: Kristina Wilfore, #ShePersisted.

  Includes sections on pre-emptive action and "How to report and document attacks online, while obtaining the necessary technical and psychological support throughout".

  https://r2g26a.n3cdn1.secureserver.net/wp-content/uploads/2022/06/ShePersisted_Digital_Resilience_Toolkit.pdf

- HateAid reporting form (2018)

  Author: HateAid

  An organization based in Germany with a self-reporting form for whether "you've experienced digital violence yourself, witnessed online attacks, or want to report cases of hate speech online".

  https://hateaid.org/en/reporting-form/

- Safety Training for Female Journalists (2021)

  Author: Free Press Unlimited (FPU)

  A "website for trainers to provide safety training to women journalists, integrating digital security, physical safety and well-being".

  https://safetyforfemalejournalists.org/

- Fix the Glitch Toolkit 2.0: Helping to End Online Gender Based Violence for Black Women (2021)

  Authors: Seyi Akiwowo, Hayle Chalke-Davies, Kiran Chalke and Layla Austin, Glitch UK

  A toolkit "designed and reviewed with experts to support Black women and those who want to help end online gender-based violence (OBGV) against Black women but may not know where to begin".

  https://glitchcharity.co.uk/wp-content/uploads/2021/04/Glitch-Toolkit-FULL-Interactive.pdf

- What to do if you are experiencing online abuse (2022)

  Author: Seyi Akiwowo, Glitch UK

  "A spreadsheet, and resource to more easily document and report online abuse."

  https://glitchcharity.co.uk/wp-content/uploads/2021/09/Glitch_Documenting_Online_Abuse_September2021.pdf

  https://glitchcharity.co.uk/wp-content/uploads/2021/05/Documenting-Online-Abuse-form.pdf

- Coalition against Online Violence (CAOV), a resource centre created by IWMF and ICFJ in 2021.

  "A hub where women journalists can come to find the latest information on online abuse and harassment, with everything in one place."

  https://onlineviolenceresponsehub.org/resources

The above link to CAOV includes the following additional resources:

- Shouting into the Void: Why Reporting Abuse to Social Media Platforms Is So Hard and How to Fix It (2023)

  Authors: Kat Lo and Viktorya Vilk

  "Resources on how social media companies can create reporting dashboards, and with a Product case study on TRFilter."

  https://pen.org/report/shouting-into-the-void/

- Games Hotline Digital Safety Guide: Protecting Yourself During Online Harassment Attacks (2023)

  Authors: Jaclyn Friedman, Anita Sarkeesian, and Renee Bracey Sherman, updated by Games and Online Harassment Hotline with Tall Poppy

  Focus on "how to keep yourself safe from individuals, loosely organized groups and cybermobs… especially designed for women, Black, indigenous, and people of color, trans people, and everyone else whose existing oppressions are made worse by digital violence", with a "suggested list of steps to mitigate potential escalation of attacks through intimidation, harassment, threats, and abuse".

  https://gameshotline.org/online-free-safety-guide/#about

- We Keep Us Safe: LGBTQ Digital Safety Guide

  GLAAD

  Includes "Common scenarios and help guides", from the leading national LGBTQ media advocacy organization in the US.

  https://glaad.org/smsi/lgbtq-digital-safety-guide/

- OntheLine - protocol for newsrooms responding to online violence

  International Press Institute (IPI)

  Many checklists available e.g., https://newsrooms-ontheline.ipi.media/measures/forms-2/

  https://newsrooms-ontheline.ipi.media/lessons/session-1-building-an-effective-protocol-initial-steps/

  https://static1.squarespace.com/static/642a4483efd0ce42e3a2c9f3/t/64938be2571403o9c22d4544/1687391202905/Documentation+-+OnlineSOS+Checklist.pdf

- Editor's checklist: Protecting staff and freelancers against online abuse (2023)

  Author: Committee to Protect Journalists

  A form which "allows editors and commissioners to understand how well-prepared journalists are when it comes to protecting themselves against online abuse".

  https://cpj.org/2022/07/editors-checklist-protecting-staff-and-freelancers-against-online-abuse/

  https://cpj.org/wp-content/uploads/2022/07/CPJ_online_abuse_checklist.pdf

- Online SOS (2023)

  A "non-profit organization connecting people with information and tools to take action in the face of online harassment"; includes a "digital security cheat sheet to review the possible accounts and places your personal and professional information might be stored", a guide for therapists supporting journalists, and a Threat Modeling form: https://static1.squarespace.com/static/642a4483efd0ce42e3a2c9f3/t/6444b6ace62dac013a8a0034/1682224812966/Threat+Modeling+-+OnlineSOS+Checklist.pdf

  https://www.onlinesos.org/

  https://www.onlinesos.org/fortherapists

- Online Harassment Field Manual (2018)

  Author: PEN America

  Has a guide for documenting online harassment, including tips for downloadable screen-capturing apps and documenting abuse via emails, and a look at the laws in the United States.

  https://onlineharassmentfieldmanual.pen.org/documenting-online-harassment/

- A Mental Health Guide for Journalists Facing Online Violence (2022)

  Author: Ana Maria Zellhuber Pérez and Juan Carlos Segarra Pérez of Vinland Solution, S.A de C.V, International Women's Media Foundation (IWMF)

  This provides a "mental health self-evaluation chart so journalists can assess how online violence is affecting their wellbeing".

  https://www.iwmf.org/mental-health-guide/

- Guide to Protecting Newsrooms and Journalists Against Online Violence (2022)

  Author: Ela Stapley, International Women's Media Foundation (IWMF)

  Policies and best practices for newsrooms, which includes a reporting and escalation policy template: https://docs.google.com/document/d/1iXr9ajt88Wp1lMQk-A5GAUMvuBqknZIi9Ft4Y8m4j-w/edit

  https://www.iwmf.org/newsroom-policy-guide/

- Vita Activa

  Provides online support and strategic solutions for women and LGBTIQ+ journalists, activists and gender, land and labour rights, and freedom of expression defenders, in Spanish and English.

  https://vita-activa.org/

- Tall Poppy: resources

  Includes links to resources for people experiencing image-based sexual abuse. A team which "helps protect from online harassment, fraud and social engineering".

  https://www.tallpoppy.com/resources

- Right To Be's storytelling platform

  "A safe space where you can share your harassment story, get support, and help others experiencing harassment"; since 2005, they have received 32,000 stories of harassment.

  https://stories.righttobe.org/

- Hamara Internet: Cyber Harassment Helpline

  A digital platform based in Pakistan with a free and confidential helpline service for anyone "being harassed, bullied, or threatened online", as well as an "online harassment quiz".

  https://hamarainternet.org/

  https://hamarainternet.org/crisis-center/crisis-center-quiz/

- Sample Technology Abuse Log (2014)

  Safety Net Project, National Network to End Domestic Violence (NNEDV), US.

  https://static1.squarespace.com/static/51dc541ce4b03ebab8c5c88c/t/59ea06528dd041e0d71d8b5a/1508509266850/Sample+Documentation+Log_2014.pdf

  https://www.techsafety.org/documentationtips

- Report an Antisemitic, Bias or Discriminatory Incident

  Anti-Defamation League (ADL), based in the US, has a mission "to stop the defamation of the Jewish people and to secure justice and fair treatment to all."

  https://www.adl.org/report-incident

- Not just words: How reputational attacks harm journalists and undermine press freedom (2023)

  Authors: Chris Tenove, Ahmed Al-Rawi, Juan Merchan, Manimugdha Sharma, and Gustavo Villela, Global Reporting Centre, in partnership with the UBC School of Journalism, Writing, and Media, the Committee to Protect Journalists, the Disinformation Project at Simon Fraser University, and PEN Canada.

  "Understanding reputational attacks against journalists, including the gender factor, and understanding newsroom protocols around them."

  https://globalreportingcentre.org/reputational-attacks/

  https://globalreportingcentre.org/reputational-attacks/report-full.pdf

- Malign Creativity: How Gender, Sex, and Lies are Weaponized Against Women Online (2021)

  Authors: Nina Jankowicz, Jillian Hunchak, Alexandra Pavliuc, Celia Davies, Shannon Piersonand Zoë Kaufmann, The Wilson Center.

  Includes "recommendations for lawmakers, technology policymakers, and social media users".

  https://www.wilsoncenter.org/publication/malign-creativity-how-gender-sex-and-lies-are-weaponized-against-women-online

- Online gendered disinformation and sexist hate speech (2023)

  Authors: Malin Palm and Nynne Storm Refsing, International Media Support

  "Learning brief focuses on the issues of online gendered disinformation and sexist hate speech against women, girls and non-binary people who work or appear in the media and what media development organisations can do to address them."

  https://www.mediasupport.org/publication/online-gendered-disinformation-and-sexist-hate-speech/

- Digital Safety Snacks

  Authors: PEN America, the Online News Association, and the International Women's Media Foundation.

  "Step-by-step videos to help you defend yourself against online abuse."

  https://pen.org/digital-safety-snacks/

- Attacks and Harassment: The Impact on Female Journalists and Their Reporting (2018)

  Author: Michelle Ferrier, Trollbusters and IWMF

  "An early supporter of women and journalists targeted by online harassment in the field."

  https://yoursosteam.wordpress.com/research-on-online-abuse/

- The Intersectionality and Cybersecurity Toolkit (2022)

  Authors: Marissa Conway and Nehmat Kaur, Centre for Feminist Foreign Policy UK

  "This toolkit aims to equip its readers with how to use an intersectional lens to explore and rethink cybersecurity."

  https://static1.squarespace.com/static/57cd7cd9d482e9784e4ccc34/t/6231aa615a838779 0df1daa5/1647422050254/The+Intersectionality+and+Cybersecurity+Toolkit.pdf

PUBLISHED BY OSCE RFOM

Wallnerstrasse 6
A-1010 Vienna, Austria

Tel.: +43-1 514 36 68 00

E-mail: pm-fom@osce.org

https://www.osce.org/fom/safety-female-journalists-online

Produced in collaboration with the International Center for Journalists (ICFJ)

ISBN 978-92-9271-240-2

SEPTEMBER 2023

This publication has been made possible thanks to financial contributions from Bulgaria, the Czech Republic, Luxembourg, Sweden and Switzerland.

"The Guidelines for Monitoring Online Violence Against Female Journalists represents the independent work and expertise of the researchers, and the views and opinions expressed are those of the authors. It does not represent the policy of views of organizations providing funding"