

ONTHELINE PROJECT

Online Harassment: Mitigating the emotional and professional impact of online hate

NewsSpectrum | March 8, 2022

JAVIER LUQUE-MARTÍNEZ

IPI Head of Digital Communications

jluque@ipi.media



4 MODULES:

- 1. Why IPI has worked on online harassment and the different phases of our research (5 minutes)
- 2. Types of attacks (5 minutes + Q&A)
- 3. Mitigating emotional impact: Building your own plan of action as a journalist (15 minutes + Q&A)
- 4. Mitigating the reputational impact: Investigating coordinated campaigns (15 minutes + Q&A)



M1: PHASES OF IPI'S RESEARCH

IDENTIFYING THE PROBLEM 2014 2015

Is this an issue for the press? We went from reporting about individual cases to carry out 2 press freedom missions: Circa 50 interviews in Turkey and Jordan. Outcome: Yes.

UNDERSTANDING THE SCOPE OF THE PROBLEM 2016 What's the scope of the problem? <u>Quantitative Approach</u>: We created a database of the attacks based on the manual analysis of over 1000 tweets in Turkey and Austria. Outcome: we spotted trends, topics, type of abuse, gender.

2017

How these online attacks operate? <u>Qualitative Approach</u>: We joined forces with other researchers from Jigsaw, IFTF, and Global Voices. Outcome: We identified the elements of coordinated attacks - hacking, phishing, hashtags, botnets, smear content -.

MEASURES FOR NEWSROOMS AND JOURNALISTS 2018 2021

What can we do about it? Research of best practices in over 40 newsrooms across Europe + Development of the protocol + Development of the training workshops with newsrooms and journalists Best practices in countering online hate.



PROTOCOL FOR NEWSROOMS





2018 – Working visits to 5 countries: Finland, Germany, Poland, Spain and the UK.

40 NEWSROOMS & JOURNALISTS ASSOC.

110 EDITORS, JOURNALISTS & SOCIAL MEDIA MANAGERS AND MANAGERS

4 FOCUS GROUPS ON FREELANCERS

4 FOCUS GROUPS ON GENDER VIOLENCE

Feb. 2019 - Expert meeting in Vienna.

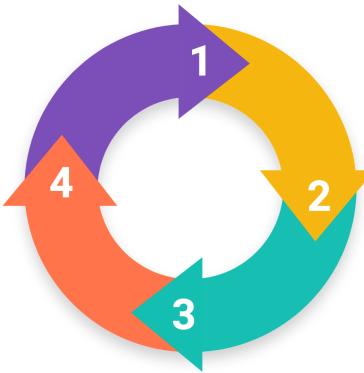
14 EXPERTS FROM ALL THESE COUNTRIES AND AUSTRIA JOINED TO HELP DEVELOP THE PROTOCOL FOR NEWSROOMS.



How does the protocol work?

1. Reporting

Create a culture of safety around online abuse and establish clear channels to report it





2. Assessment

Evaluate the risk for the targeted journalist and media outlet



3. Support

Implement support and response measures to limit impact



4. Reassessment

Track the evolution of cases and the efficiency of support measures

M2: TYPES OF ONLINE ATTACKS

PRELIMINARY NOTES

- We have to discern between: attacks by individuals and coordinated attacks / organised campaigns. Different objectives + some distinctive elements.
- Gender-based violence: Women journalists are disproportionately more attacked than men, and the viciousness or violence of the online abuse directed at women is far way worse than the one targeting men journalists.
- A tool of a set of tactics to control the narrative: Online abuse, specially organised campaigns, is part of a broader effort of propaganda to corner quality, independent journalism.



M2: TYPES OF ONLINE ATTACKS

ATTACKS BY INDIVIDUALS

○ Threats of violence

- Death threats, threats of physical violence, threats of rape and other direct forms of violent intimidation, such as threats directed at relatives (children, partner, etc.) – The latter is a type of attack very distinctive on women journalists
- Implied threats (70 percent), including doxing. SM, comments, emails, etc. Goal: Intimidate, silence + Impact: Both professional (self-censorship) and emotional impact (long exposition might result in PTSD symptomology).

○ Abusive behaviour

 Insults, physical appearance, drop-drop-drop. "Trolling". SM, comments, emails, etc. Goal: Intimidate, silence. Impact: Emotional

○ Smear attacks

 Doctored graphs (memes), pictures of the journalists, hit pieces. SM, comments, emails + propaganda sites. Goal: Discredit the message by discrediting the messenger + Impact: Both emotional and professional impact.



M2: TYPES OF ONLINE ATTACKS

ORGANISED ONLINE ATTACKS

 Message might contain a hashtag, labelling the journalist

First wave usually lasts 24 to 48 hours

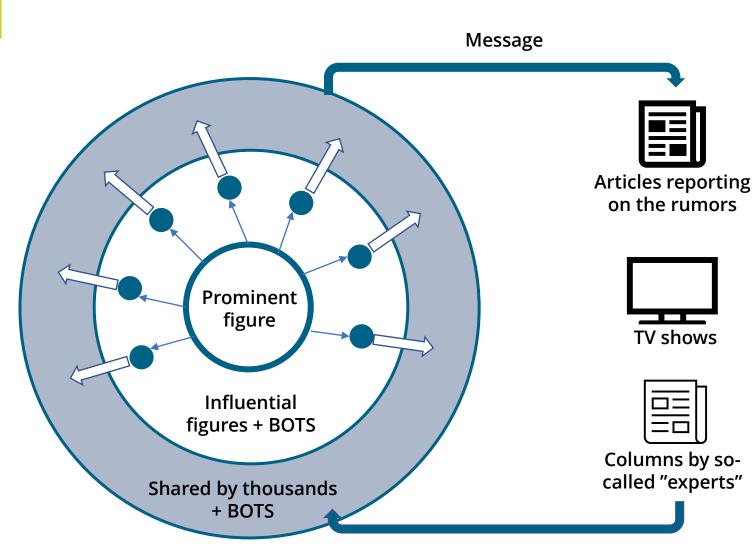
○ From then on, the same journalist will be labelled and targeted with the same message

GOAL

○ Create a suffocating atmosphere

 Give the impression that the rumor that is shared by a few thousands is what most of the society thinks

 Discredit the messenger + bring her to silence





MEASURES

COMBINED APPROACH

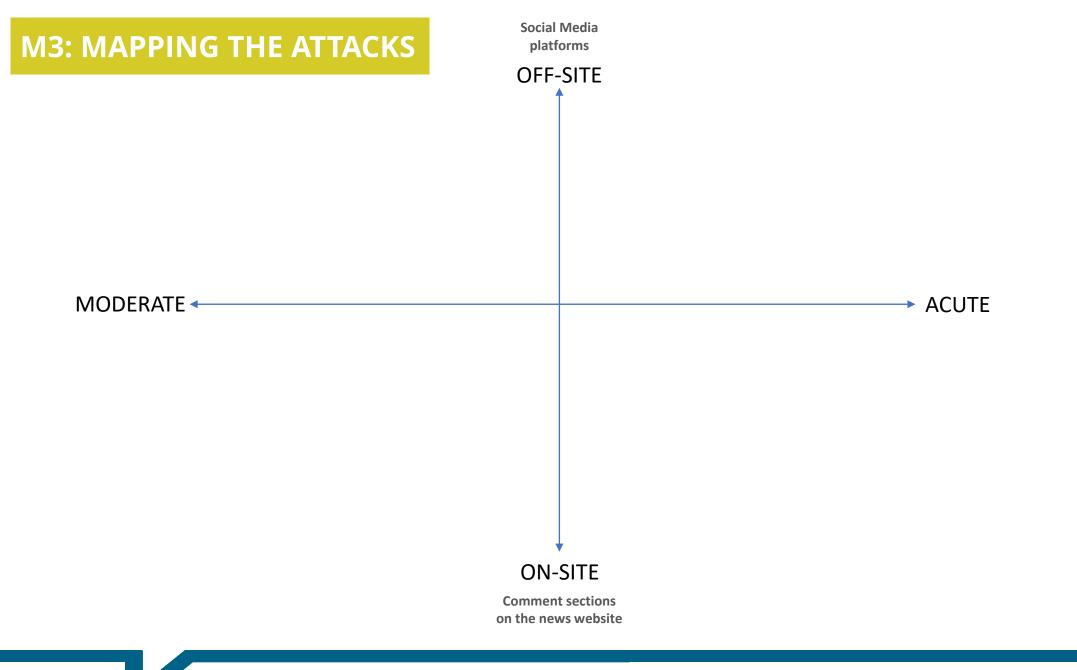
Mitigate the emotional impact – mitigate potential self-censorship.

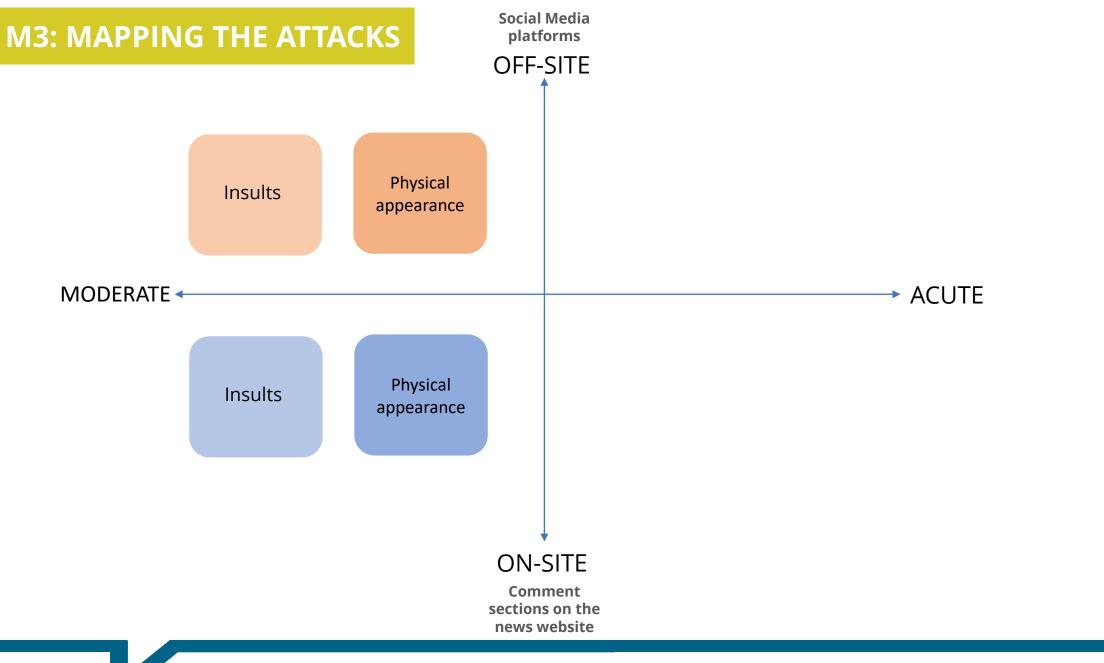
- Taking control over the attacks Mapping the attacks:
- Develop your own protocol. Use it whenever you feel like it.

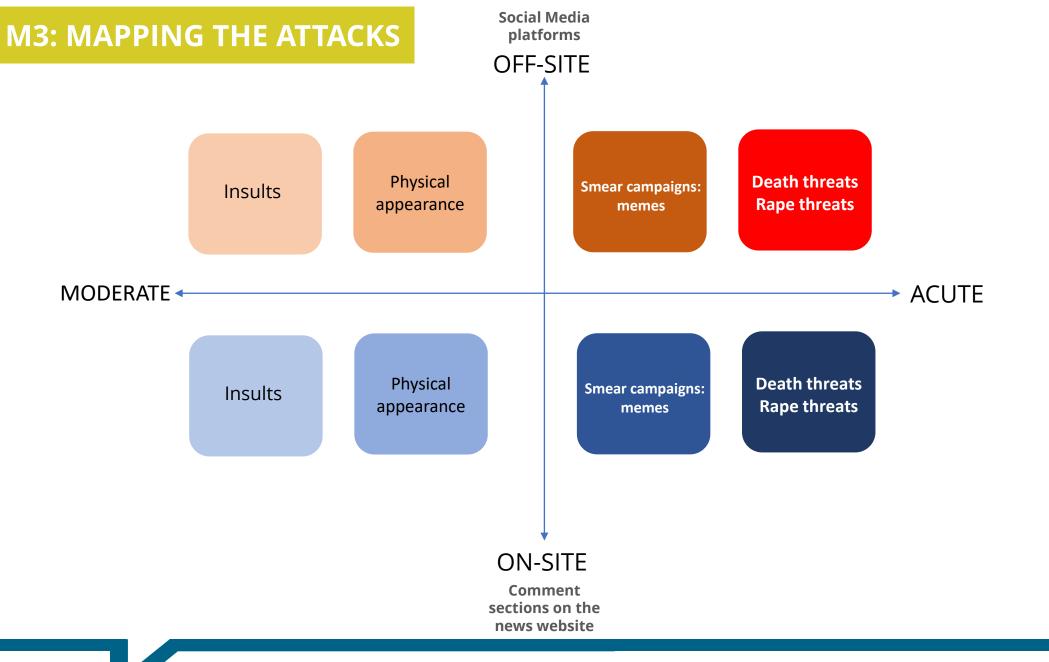
○ Mitigate the reputational

- Exposing disinformation, smear campaigns targeting your audience. Because we are journalists – Turn the tools of investigative journalism back onto the problem: Who? What? Why? How? When? Where?...
- Reduce the emotional impact on the (individual) targets: The attack is perceived as an object of study instead of a personal attack.
- Because it increases the revenue through, e.g. growing subscriptions.









PREVENTION

- Disable any geolocation features on social media. Avoid posting messages on social media that make your daily routines obvious.
- Revisit your social media accounts and consider removing photographs with identifiable locations that could reveal your patterns or personal information (anything from identifying surroundings to license plate numbers). Do the same for information about family members on your accounts: Women journalists in particular are prone to receiving comments directed at their family members.
- Disable permission for tagging or mentioning your account on social media.
- If you are doing confidential work, use encrypted channels to communicate with sources.
- Where possible, take part in formal and informal peer support mechanisms within the media organization or colleagues. Peer support networks have proven to be an important tool to minimize the emotional toll on those who are targeted with online violence.



REACTION

○ In the event of immediate danger or physical threat:

- If you are away from the newsroom, contact the police immediately, followed by your editor, the online safety expert and any relevant peer support groups.
- Consider asking another person to take screenshots to document all the threats or intimidating posts as way to limit your exposure to the material.
- If you are in the newsroom, immediately contact your editor and stay in the newsroom until adequate measures have been put in place and contact your family.
- O <u>https://newsrooms-ontheline.ipi.media/freelance-journalist/</u>



REACTION

In case you are the target of a smear campaign or wave of hate consider taking the following steps:

- Document the abuse (e.g., take screenshots) or ask a colleague to do so. Avoiding excessive exposure to these kinds of posts is a good strategy to minimize further emotional and professional impact.
- Silence the whole conversation especially in the first 24-48h, but have someone who can monitor it from another account in case there are more credible threats.
- Carefully consider if you want to respond to any of the abusive comments. Generally speaking, the best course of action is to avoid engaging with users who have targeted you. Most experts have found that doing so does not lead to an improvement and may worsen the abuse.
- If you do choose to respond, it is recommended to use extreme politeness, irony and humor as ways to diffuse the vitriol.
- If it is not you but a colleague who is targeted, coordinate with other journalists to engage in online counter speech and show public support.



REACTION

In case you are the target of a smear campaign or wave of hate consider taking the following steps:

- Hacking or phishing attempts? Contact a digital security expert through the journalist association or a colleague.
- Legal threats? Consider contacting your newsroom or journalist association for a first assessment.
- Emotional support: Consider contacting your newsroom or journalist association for a first assessment

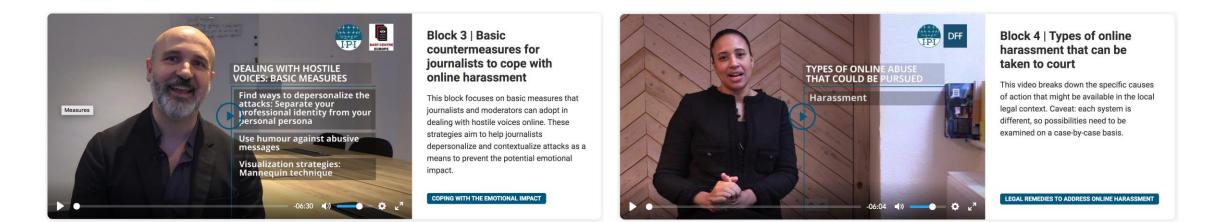


SUPPORT MECHANISMS

• The aim of all these mechanisms is to ensure that targeted journalists can do their jobs safely.

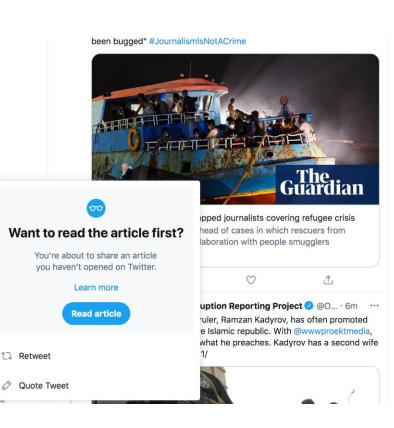
○ Legal and Psychological support: Video Tutorials.

- Coping with the emotional impact
- Legal remedies to address online harassment





AUDIENCE MODERATION



Moderating online abuse ON SOCIAL MEDIA PLATFORMS

O Twitter

- Muting: When it comes to online abuse, moderators tend to mute rather than block accounts. This
 option dilutes the direct impact of the abuse as the target will no longer receive notifications from
 the muted account. It also prevents a possible angry backlash as the muted user has no knowledge
 of the muting. Finally, muting allows moderators to still see content produced by muted accounts
 and therefore remain vigilant to any potential credible threats against the media outlet or a
 journalist.
- Blocking: Moderators tend to block accounts that persistently spam or send scams.
- **Reporting:** Moderators generally report tweets or accounts to Twitter that disseminate potentially credible and imminent threats or contain violent imagery Important for legal purposes.
- Hide replies: Tweet authors have the option to hide replies to their Tweets. Everyone can still access hidden replies through the hidden reply icon, which shows up on the original Tweet when there are hidden replies. Additionally, the Tweet author can unhide a reply at any time. When a Tweet author hides a reply, the author of the reply will not be notified.
- Select who can Reply to your tweets: Everyone, Accounts that YOU follow or accounts that are Tagged in your tweet. + "Silence" the whole conversation, so you don't receive notificiations.



Moderating online abuse ON SOCIAL MEDIA PLATFORMS

Facebook

- Delete a comment when it contains aggressive or threatening content or derogatory words and insults. This is done to promote a healthy public discussion. Criticism, no matter how harsh, is permitted, however.
- Hide a comment with abusive content. Moderators generally considered this less effective than deleting as the user and the user's friends can still see content in question, even if others cannot.
- Ban a user from the media outlet's Facebook Page when the user has repeatedly posted hateful or abusive comments, even after being warned by the moderators. This is done to remove a user who is seen as persistently undermining the values of a health discussion and the open community that the media organization aims to generate.
- Remove a user from the page, as a warning to deter further abusive comments. Less consequential than banning as the user can like or follow the Page again.
- Disable/turn off comments, although this feature is only available on video posts. This is one when the moderation team does not have the resources to moderate the flow of comments on a video or live stream.
- Block words and set the strength of the profanity filter
- Report a post or a Page that the moderator considers has breached both Facebook's and the media's own community standards. Facebook's stated policy is that it will then assess whether the content or Page should be removed.

ne ++++ +++++/ !·P·I

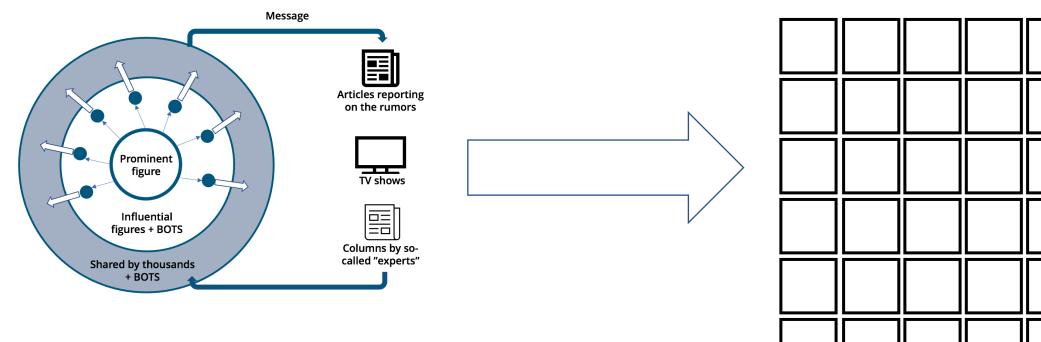
AUDIENCE MODERATION

ASSESSMENT	TYPE OF HARASSMENT	ACTION	MODERATION	DOCUMENTATION	COMMENTS
CHRONIC / OFF-SITE	Insults	NONE	TWITTER: REPORT IT MUTE / BLOCK IT FACEBOOK / IG: REPORT IT HIDE THE COMMENT AS WELL	DON'T	I WILL ASK COLLEAGUES TO TAKE OVER MY SOCIAL MEDIA ACCOUNTS ONLY IF THERE'S AN AVALANCHE OF INSULTS. I JUST DON'T FEEL LIKE DEALING WITH EVERY SINGLE ONE OF THE INSULTS
	Physical appearance	 RULE OF THUMB: Do not engage, but if I decide to engage, remember: Use irony, humor. Don't show it affects me. Name and shame – Indirectly fosters positive reactions from my followers 	TWITTER: REPORT IT MUTE / BLOCK IT FACEBOOK / IG: REPORT IT HIDE THE COMMENT AS WELL	 Document the abuse(e.g., take screenshots, name, URL) or ask a colleague to do so. If I choose to report the attack, document the confirmation of report 	I WILL ASK COLLEAGUES TO TAKE OVER MY SOCIAL MEDIA ACCOUNTS ONLY IF THERE'S AN AVALANCHE OF INSULTS. I JUST DON'T FEEL LIKE DEALING WITH EVERY SINGLE ONE OF THE INSULTS
ACUTE / OFF-SITE	Smear campaigns	 Coordinate with other journalists to engage in online counter speech and show public support. Name and shame – Indirectly fosters positive reactions from my followers. Ask your colleague to perform a quick doxing exercise of your digital footprint to understand if there's more information about you out there that can be used against you. 	TWITTER: REPORT IT MUTE FACEBOOK / IG: REPORT IT HIDE THE COMMENT AS WELL	 Document the abuse (e.g., take screenshots, name, URL) or ask a colleague to do so. Document the confirmation of report Take screenshot & Document smear articles. 	I WILL ASK COLLEAGUES TO TAKE OVER MY SOCIAL MEDIA ACCOUNTS ONLY IF THERE'S AN AVALANCHE OF INSULTS.
	Death threats Rape threats	 Away from the newsroom, contact the police immediately + Then my editor or colleague: Telf Police: +123456 / Telf. Colleague/editor: +98765 In the newsroom, immediately contact your editor and stay in the newsroom until adequate measures have been put in place and contact your family. As your colleague to anakyse if there's more personal details or those of your closest relatives that can be used to threaten you. 		 Document the abuse (e.g., take screenshots, name, URL) or ask a colleague to do so. Document the confirmation of report 	I WILL ASK COLLEAGUES TO TAKE OVER MY SOCIAL MEDIA ACCOUNTS TO MONITOR FURTHER CREDIBLE THREATS.



GOAL 2: PLANT SEED OF THE DOUBT IN YOUR AUDIENCE - QUALITY JOURNALISM

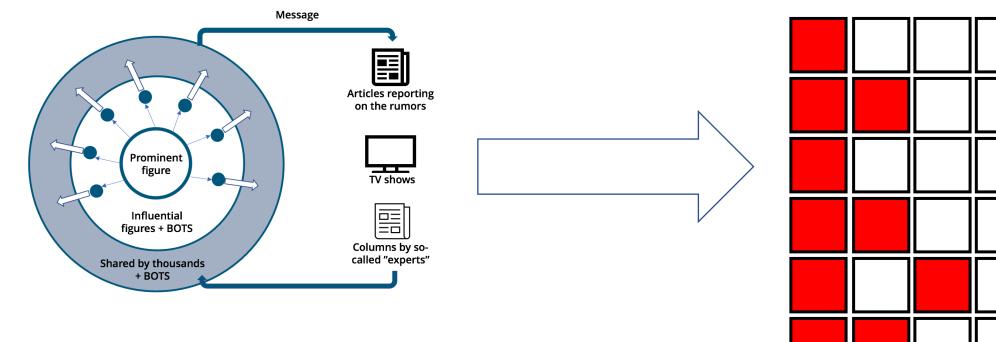






GOAL 2: PLANT SEED OF THE DOUBT IN YOUR AUDIENCE - QUALITY JOURNALISM







M4: ORGANISED ONLINE ATTACKS

How?





JAVIER LUQUE-MARTINEZ, IPI Head of Digital Communications | March 8, 2022 | NewsSpectrum

EXPOSING DISINFO/ SMEAR CAMPAIGNS

EXPOSING DISINFO/ SMEAR CAMPAIGNS



RAPPLER



Oxford Reuters Institute, Research, 2019.

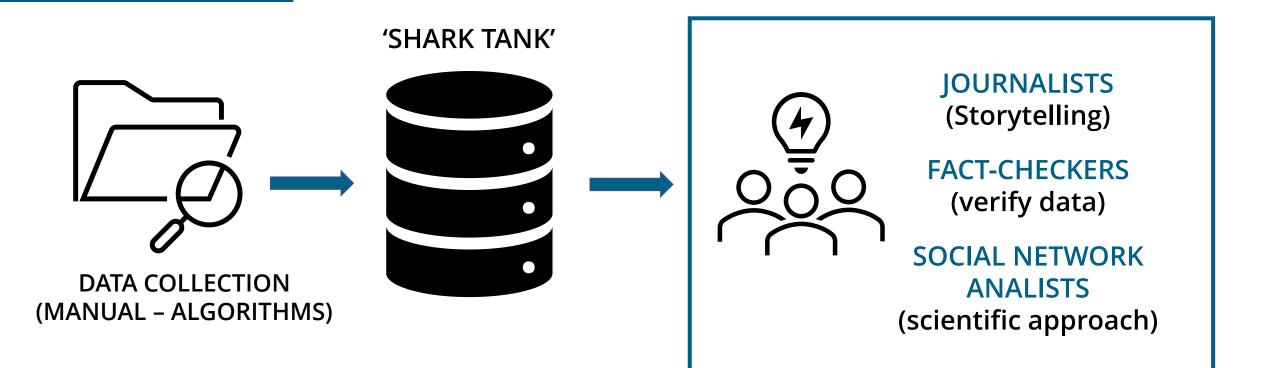
○ Turning point: May 2016

- Rodrigo Duterte wins the Presidential elections in The Philippines and kicks off the so-called "War on Drugs".
- According to Maria Ressa, within weeks they saw how disinformation and smear targeting critics of Duterte's "War on Drugs" were increasingly disseminated on social media platforms, especially Facebook, where most of Rappler's operations were running at that moment.
- 2017 Duterte's State of the Nation Speech: Targets specifically Rappler and Maria Ressa, labelling them as: 'Fake News (Trumpism)', 'Elitists (not relatable to people)' and 'fully American owned' (foreign agents – Russia now).
- Almost immediately after: Maria Ressa and all Rappler's national and regional reporters were banned from having direct access to Duterte or the Presidential Palace + the wave of online attacks increased significantly + legal harassment - Ressa currently faces up to 100 years in prison if convicted in a series of alleged violations of tax, foreign ownership and cyber libel.





EXPOSING DISINFO/ SMEAR CAMPAIGNS







M4: ORGANISED ONLINE ATTACKS





OTHER RESOURCES

THE BIG PICTURE

- In general, it is useful to create some distance from your specific case. Try to view online harassment as a larger societal phenomenon rather than abuse directed at you personally.
- Understand that online aggressors are attacking you because of your work – even if the attacks refer to your personal life. This can be helpful in easing the emotional and personal burden.
- Consider putting your journalistic skills to work as a response to online harassment. We know that in many cases harassment is not organic but is planned or at least driven by a network of users linked to political parties or governments. Revealing these networks can help deprive them of their power and influence.





ONTHELINE PROJECT

THANK YOU!

STAY TUNED FOR OUR UPCOMING ONLINE WORKSHOPS BOTH FOR NEWSROOMS AND JOURNALISTS

JAVIER LUQUE-MARTÍNEZ

IPI Head of Digital Communications

jluque@ipi.media