

Responding to the challenges caused by information incidents

26 November 2020

Purpose

We know that certain events can affect the information environment by prompting an increasing complexity of accurate information, confusion, or by creating information gaps - all of which can result in an increase in the volume of misinformation. This was clearly evident during the coronavirus pandemic, which prompted a slew of intensified counter-misinformation measures from internet companies, governments, media, fact checkers, academics and civil society.

The response to coronavirus misinformation this year has shown how fast and innovatively those working to analyse and counter it can respond. But it has also thrown light on the need for greater discussion of principles, proportionality, and the use of evidence in responding to other types of future information incidents that may be just around the corner.

That is why Full Fact is bringing together practitioners, experts and community groups from different sectors affected by and aiming to affect the information environment to develop a framework to identify the issues that occur during moments of crisis and develop joint aims for how organisations should respond. Our aim is to develop something simple and useful that can help specialists in this area coordinate our work, and outside stakeholders understand it.

Outline

Paper 1, published on 20 November, outlined the types of incidents a framework could be applied to, and set out an initial structure for assessing whether incidents pose risks over and above day-to-day misinformation and therefore could merit an exceptional, collaborative response.

In this paper we outline an overview of the emerging framework components, and present a set of challenges that we judge are likely to arise within an individual incident. For each challenge, we offer a corresponding high level aim for responding, alongside some practical real-world examples of such responses.

Our goal is to produce a simple and useful framework that helps assess, respond to, and demonstrate success in tackling information incidents – whilst recognising the deep complexity of information environments, and the nuance and care needed to respond to incidents which affect how information is produced, disseminated and consumed.

Annex A sets out the detail of key sections of the framework. Below we give an overview of these three categories.

Overview of emerging framework

Challenges

Full Fact identified nine high level challenges by analysing common characteristics of incidents with which we are familiar, as a UK-based organisation. We then sought input from colleagues working in Canada, India, Nigeria, Spain, Argentina and the USA. As this project progresses we will work with experts, practitioners and communities with experience of different information environments to build on this – for example, how to better describe the adversarial and industrialised nature of disinformation in some countries, or factoring in motives for creating and distributing bad information.

Aims

One of the benefits that we hope users of this framework find is in having aims that different types of organisations use to centre their responses around. Having a shared terminology supports greater communication, by making it simpler to understand one another during times of high pressure. By also having a joint understanding of what the common goal is, we hope that organisations can better discern why certain interventions are taken and greater collaboration is supported.

Therefore these aims are targeted at those who will be making decisions on how to respond to incidents, as well as those who wish to scrutinise or evaluate those responses. The aims are not mutually exclusive and should be seen as complementary.

We understand that not all of these aims will be the priority for each organisation. However, the more we understand each other's objectives in actions, the easier it will be to have joint and/or complementary efforts. As with the challenges, the aims presented below are our initial assessment and will be refined with further feedback and analysis.

Practical responses

Some actions are clearly better delivered by one organisation or sector than others. For example internet platforms will have insight into what's trending on their platforms, while governments are the only actor which can effectively plan for the long term quality of a country's official information provision. Meanwhile, fact checkers and academics can provide credible, independent analysis and assessment of businesses' and governments' actions. This framework aims to recognise and acknowledge these strengths, and create a structure for all those countering misinformation to bring our experience, knowledge and resources to a collective response.

The suggested responses set out here are in no way exhaustive. They hope to inspire rather than direct. Wherever possible, it is suggested that the responses chosen to fulfil the aims should seek to be:

- Credible, particularly to the target audience as well as independent researchers
- Efficient, in tackling a clearly defined problem or targeting a clearly defined audience
- Scalable, which could mean ability to respond at pace or to large volumes

It is inevitable that there will be difficulty in designing interventions which successfully tackle all three criteria, but we hope these are helpful framing points. Independent evaluation will be key to truly understanding the impact of interventions and ensuring that they are effective and proportionate.

Next steps

We are publishing this paper to encourage wider feedback on how our thinking is developing. We would be interested to hear thoughts from other organisations who are involved in responding to bad information on:

- Thinking about efforts to combat misinformation in exceptional circumstances, do the challenges described correspond to your experience? Is there anything missing?
- Are there other examples of these challenges which could help different actors involved in tackling misinformation incidents to better understand each other and work together?
- Looking at the high level aims in bold in the right hand column of the table below, are there any missing which should be included here?

The third paper will explore what capabilities are needed for a framework like this to function, including resources, roles, information sharing, evaluation. The first draft of a framework will be published in early 2021.

Please do get in touch if you have any feedback on this paper, or would like to contribute to this work, at phoebe@fullfact.org. Please note we may be unable to respond to every contribution.

Annex A: Identifying aims for responding to challenges

Challenges

Aims and example responses

1. Threats to freedom of expression e.g. when there is:

- Lack of independent scrutiny of laws, moderation policies and norms that allows for censorship creep
- Unprecedented use of technology to reach large audiences without the ability for independent scrutiny
- Suspected or known foreign interference

Design responses that are demonstrably proportionate to clearly identified harms, and open to informed debate and discussion

For example:

- Provide access to engagement, trends, and advertiser data to enable independent research on the impact of responses
- Evaluate effectiveness of counter-misinformation efforts and publish learnings
- Enable independent experts to scrutinise AI recommendations

2. Unclear or quickly changing situation e.g. when there is:

- Lack of insight into type and scope of misinformation and/or movement of content between platforms
- Unhelpful duplication of efforts among organisations
- Contradictory interpretations of a situation

Work towards a shared assessment of the situation and complimentary responses

- Share monitoring and verification information between trusted experts
- Support smaller platforms to share trends data to help predict when narratives / claims might move to mainstream platforms
- Brief media and other mainstream sources of information to reduce risk of amplification and stop dissemination of harmful information

3. Difficulty disseminating or communicating information e.g. when there is:

- Low baseline knowledge of key issues among public, politicians and media
- Low statistical literacy among public and media
- Accurate information is not contextualised or adapted for certain groups
- Topics are complicated or highly technical
- Information overload and audiences find it hard to judge content in the decontextualised format of news feeds
- Intense partisanship / emotive topics make it harder for accurate information to be believed

Ensure good information reaches both affected groups and the wider public, and the key information is communicated effectively by trusted figures

- Promote relevant impartial or official sources of information
- Identify and engage with appropriate trusted voices to disseminate information
- Disseminate information to pre-empt belief in emerging conspiracy theories

4. Information vacuums and uncertainty e.g. when:

- Information is partial, allowing for distorted reporting and discussion
- New information must be produced, leaving a temporary gap
- Official advice is changing quickly or official sources backtrack
- Future is unknown so unfounded claims of certainty gain traction

Ensure reliable information from authoritative sources is available and that any limitations are communicated

- Funding and resources for statistical offices and impartial information providers
- Horizon scanning to ensure information is adequate for future public decisions
- Transparently explain why information or advice has changed
- Strengthen and support impartial journalism

5. Unhelpful behaviour by influential public figures e.g. when:

- High profile figures repeat false claims or make conflicting statements
- High profile figures cast doubt on accurate information
- High profile figures deliberately encourage distrust of mainstream media

Provide context to help audience make judgements and promote alternative trustworthy sources of information

- Apply warnings, pop-ups and labels
- Promote alternative coverage from trustworthy media and fact checkers
- Give information and caveats about sources of information being presented

6. Pressure to work at speed and scale to halt spread of false beliefs e.g. when:

- Volume and speed of information increases beyond resources of human teams to monitor and counteract it
- Increased consumption of news encourages media to report insignificant stories as major developments and increases likelihood of mistakes being made
- Unintended consequences arise from responses including entrenchment of false beliefs

Limit bad information and ensure corrective information appears when and where people need it, and have a clear plan for scaling

- Remove or suppress content and/or persistent offenders in a proportionate manner
- Design effective corrective content
- Implement additional verification standards before information is disseminated
- Strengthen moderation enforcement policies
- Invest in burst capacity and systems including support for experts and news organisations
- Work with volunteers to feed AI with marked up data for emerging topics or claims

7. Immediate threats to public order and safety e.g. when:

- Public order and safety is dependent on the public understanding information accurately
- Communication from affected communities and first responders is compromised or ignored
- False information creates potential for physical harm through violence or hazard

Consider targeted measures for affected audiences to see and trust accurate information

- Adapt or contextualised information to reach target / affected audiences
- Identify and engage with appropriate trusted voices to disseminate information

8. Lasting longer term impacts of an incident or incidents e.g. when:

- The incident spawns or entrenches conspiracy theories or myths which outlast the incident
- False narratives are repeated over years and create hard-to-shift public misperceptions

Build audience resilience, and communicate and debunk effectively

- Cross sector investment in effective communication of information
- Increase audience awareness of and ability to identify bad information
- Research and fund effective teaching methods for information literacy, and evaluate existing information literacy programmes
- Work with schools, universities and qualifications bodies to ensure critical thinking and information literacy curriculums are effective and regularly evaluated and updated