# Towards a framework for information incidents

## Paper 3: Levels of Incidents

**18 December 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 in 2020 shows 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 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. The intended result is to develop a simple and useful framework that can help specialists in this area coordinate their work, and be accessible and valuable for all stakeholders.

## Outline

This paper expands on the indicators to determine risk outlined in paper one, to identify five levels of severity that an incident could be classified as. In this paper we highlight the work that we have drawn on to develop the five levels, provide detailed

draft descriptions of each of the levels and outline four case studies to illustrate how we would classify different situations in recent history.

## Identifying levels

Throughout this process we have sought to build on and consolidate research and analysis from a range of key experts in recent years, some of whom have provided feedback on this framework. We are grateful to the following reports for laying the groundwork for our analysis:

- Carnegie Endowment for International Peace's **ABCDE framework**, which breaks down the disinformation problem into smaller components, covering the actor, behaviour, content, degree of severity or harm, and effect.
- First Draft's **7 Types of Mis- and Disinformation**, which breaks down the information ecosystem by type of content being created and shared, motivations of those who create content, and the way the content is being disseminated.
- Claire Wardle and Hossein Derakhshan's **Mis/Dis/Mal definitions** for their Council of Europe Report on information disorder. Also from this report, Wardle and Derakhshan's **Phases and Elements of Information Disorder** table which considers the agents that created, produced and distributed the example, and their motivation, the type, format and characteristics of the message; and who received and interpreted the message, and what action they took.
- Ben Nimmo's **Breakout Scale** for measuring the impact of information operations, beginning with a breakout spread within one community or platform, and culminating in a breakout triggering a policy response or call for violence.
- Joan Donovan's **Media Manipulation Lifecycle** which enables analysis of the order, scale and scope of manipulation campaigns, through five points of action charting the tactics of media manipulators.

From reviewing these reports and more, we have created a scale with five increasing levels of severity. We hope that this is a useful framework to understand the amount and type of resource that should be put into tackling an incident, to picture how incidents could evolve over time and to assess how an incident compares to others. We outline these levels below.

# Five levels of severity

## Level 1

**Business as normal; no additional response needed**

Misinformation is circulating at levels considered to be “normal”, with no incident on the horizon.
There may be some spikes around certain events, topics or locations but they die down quickly.
There is space to work on long term goals such as audience resilience and media literacy.
Key organisations are collaborating to identify emerging threats.

## Level 2

**Monitor and prepare external-facing responses**

False claims or narratives are breaking out related to a certain topic or event, on one or more smaller platforms.
The subject matter may be controversial or polarising, or has potential to threaten vulnerable groups if the volume of misinformation or engagement grows.
There is time to put plans in place to mitigate the growth and effects of misinformation.

## Level 3

**An incident is occurring; internal responses ramp up**

Misinformation is moving into the mainstream and there may be amplification (whether organic or coordinated).
The incident may be relatively short term, so acting immediately is important.
Individual organisations are implementing internal incident response plans, collaboration is enhanced to identify cross-media and cross-platform risks.

## Level 4

**An incident is occurring; coordinated responses increase**

Misinformation is moving into the mainstream and being amplified by large accounts.
There may be evidence of coordination or malicious intent.
The misinformation might spark violence or physical danger.
The incident may compound existing conspiracy theories and false narratives.
The event or topics may be completely new, or emotive, to the audience, increasing the likelihood of confusion.
Core and expanded (e.g. public health, agriculture) organisations are collaborating.

## Level 5

**Maximum response levels and co-operation required**

The incident is unusual in terms of extremely fast spread and high volumes of misinformation, and is not likely to be resolved in the short term.
The misinformation is causing, or is likely to cause, significant human harm.
Lasting effects on public misperceptions may occur.
The incident is global or affecting multiple regions, with the same misinformation often appearing in different languages.
A maximum level of cooperation from a range of different organisations is required.

We have chosen to start with a Level 1 “business as usual” with an understanding that some misinformation is, and will remain, a constant. A world in which there is no misinformation circulating is also likely to be a world with an unreasonable and dangerous amount of surveillance and censorship, and we do not advocate for that scenario. Instead Level 1 reflects a realistic scenario where there are low levels of misinformation but key organisations have capacity to focus on long term goals and priorities.

Level 5 should be reserved for only the most severe of incidents. The only example in recent history that we would classify as a Level 5 incident is the beginning of the 2020 coronavirus pandemic (see case study below).

A challenge we found in creating these levels was defining core characteristics to differentiate between Level 3 and Level 4. We have chosen to go with the length of time an incident seems likely to last for, acknowledging that responses will differ depending on whether an incident is likely to be a few days or a few months. As with all of this work, we welcome any feedback or suggestions on whether that delineation is sufficiently clear and practical.

We envision that incidents could move between levels over time, whether that is in response to rising severity, for example if it becomes clear that actually longer-term responses are needed, or where severity is decreasing, such as when an incident is drawing to a close. It may not always be clear how long an incident will last in the early stages: our aim is that this framework offers flexibility to adapt to this reality.

**Deciding on a level**

The question of who should decide what level an incident is at is one that we have carefully considered. One objective of the work to have a shared framework is to improve coordination, joint language and cohesive approaches. From that perspective it would be beneficial for different organisations to agree together what level an incident is at. However we recognise the significant concerns that come with one or multiple organisations imposing this decision on others. We also note that there are advantages to having flexibility that allows for each organisation to reach a determination of their own that reflects their experience and circumstances (this may or may not be informed by dialogue with peer entities and others, at least in the lower levels). A framework with too much choice may become inoperable: while it is difficult for multinational companies and governments to commit to something too rigid, something too loose loses its utility in enabling quick responses.

We do not at this stage recommend one approach over the other, and would welcome views.

# Case studies

Below we outline four case studies to illustrate the levels:

**Level 2: 5G conspiracies, April 2019**

- Emergence of new narratives about harm caused by new 5G technology. Potential for real world harm in the future, but no immediate threat.
- Low level of knowledge in the public and media.
- Some risk of polarisation around national identity.

**Level 3: Notre Dame fire, France 2019**

- Unexpected event, and responses to any misinformation are required immediately.
- Marginalised groups more likely to be targeted with misinformation.
- High volumes of news coverage and information (including false information) circulating.

**Level 4: Brexit Referendum, UK 2016**

- Unfamiliar event (a UK referendum) and topics (e.g. EU law and trade).
- Polarising campaign with high profile misleading claims spread by politicians.
- Unprecedented use of technology to distribute unscrutinised claims to certain groups.

**Level 5: Covid-19 pandemic, global 2020**

- Spread of high volumes of life threatening misinformation on multiple platforms.
- Extended time period and likely to continue for months longer, with lasting effects on public misperceptions and public health.
- New policies, products, partnerships and information-sharing required to effectively combat the incident.

## Once you have a level

Once a sense of the incident's severity has been established, the next step is to consult the framework which will offer direction on how to tackle the challenges which are occurring, in ways that are proportionate to severity. This is covered in our second paper, which looks at what challenges are likely to arise within an individual incident, and how organisations can respond individually, and in concert with other organisations.

## 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:

- Are these five levels of severity clear to you?
- Are these applicable to real examples you have experience of dealing with?
- What types of responses would you expect to take in each level? How would you expect the response to levels differ, for example from Level 2 to Level 4?

The first draft of the framework for misinformation incidents and crises 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.