



# TOWARDS AN IMPACT-RISK INDEX OF DISINFORMATION: MEASURING THE VIRALITY AND ENGAGEMENT OF SINGLE HOAXES

EU DisinfoLab – June 10, 2022

EU DisinfoLab's impact-risk index offers an approach to assess the potential impact of single hoaxes. The method goes through a list of eight indicators related to the virality and engagement of a single disinformative content.



Publication Date: June 10, 2022

Authors: Raquel Miguel,  
EU DisinfoLab

---

### **EU DisinfoLab**

Chaussée de Charleroi, 79  
1060 Brussels, Belgium  
E-mail: [info@disinfo.eu](mailto:info@disinfo.eu)

### **About EU DisinfoLab**

EU DisinfoLab is an independent non-profit research organisation specialised in analysing disinformation. We uncover and expose sophisticated disinformation campaigns. We seek to amplify the voices of our community of counter-disinformation experts across the EU and contribute with collective expertise to policy making. You can find more information about our work on our website

<https://www.disinfo.eu/>.

# TABLE OF CONTENTS

- INTRODUCTION: WHY AN IMPACT-RISK INDEX FOR SINGLE HOAXES ..... 3
- MEASURING INDICATORS AND SCORES ..... 4
- 1. Engagement on social networks: measuring shares and reactions ..... 4
- 2. Exposure on social networks: measuring the views ..... 4
- 3. Content circulation: the number of platforms ..... 4
- 4. Diffusion across communities: language as a proxy indicator ..... 5
- 5. Media outreach: reaching mainstream media ..... 5
- 6. Type of actor: public figures and persistent disinformative transmitters ..... 5
- 7. One narrative, many styles: appearance in various formats ..... 6
- 8. Call for action and danger of the narrative: a multiplier effect ..... 6
- THE FINAL SCORE & THE SCALE ..... 7
- INDEX APPLICATION OF FACT-CHECKED HOAXES ..... 8
- Example 1 ..... 9
- Example 2 ..... 10
- Example 3 ..... 11
- Example 4 ..... 12

# INTRODUCTION: WHY AN IMPACT-RISK INDEX FOR SINGLE HOAXES?

When confronted with a hoax, disinformation researchers often wonder about its relevance: is it just an isolated post? Will it go viral? And if it does, could it have any impact in the offline world? The study of the impact of mis- and disinformation is one of the most challenging aspects for researchers, but it is also extremely relevant.

Experts have already worked in this field with compelling results, such as Ben Nimmo, who created the [Breakout Scale](#), designed to measure the impact of influence operations. However, applying the scale to a single hoax is complex and often unfeasible. This issue inspired EU DisinfoLab to elaborate a simpler alternative that researchers can use in their daily work to analyse single pieces of disinformation. They can apply it hoax by hoax, optimising efforts to rapidly identify those falsities that potentially carry greater risk or may lead to more articulated campaigns.

In addition to this, the Digital Services Act (DSA) sets obligations for how digital services deal with content that has societal risks, and therefore measuring the impact of disinformation could be decisive in the assessment of these risks.

The **EU DisinfoLab impact-risk index** offers an approach to assess the potential impact of single hoaxes. This index could also go beyond the assessment of single hoaxes and could be applied to other claims. The method goes through a simple list of eight indicators related to the virality and engagement of a single disinformative content, whose scores will be translated into a final scale measuring the low, medium, high, or alarming impact-risk. The scale might be improved in the future to better meet the needs to the community, but it currently offers an immediate and unique method of assessment whose benefits surely surpass the possible limitations.

# MEASURING INDICATORS AND SCORES

## 1. Engagement on social networks: measuring shares and reactions

This measure takes into consideration the social media platform where the single piece of disinformation collected the highest number of **shares and reactions**, regardless of its appearance in more than one platform (which will be considered later). In fact, reactions and shares imply that there has been user's active **engagement** with the content.

- 0 – 1.000 shares and reactions = 0 points
- 1.001 - 10.000 shares and reactions = 1 point
- 10.001 – 100.000 shares and reactions = 2 points
- More than 100.001 shares and reactions = 3 points

## 2. Exposure on social networks: measuring the views

This category measures the **exposure** that users had to the content in the form of views, although we acknowledge the impossibility to fully assess whether they intentionally decided to play a video, or it automatically played. Therefore, this category, which is to be counted only for the platform gathering the highest metrics, completes the previous one.

- 0 – 1.000 views = 0 points
- 1.001 - 10.000 views = 1 point
- More than 10.001 views = 2 points

## 3. Content circulation: the number of platforms

This indicator is intended to measure cross-platform diffusion, giving one extra point if the content has been shared on more than two platforms, including messaging chat services such as Telegram (due to the social network similar function of the channels) and WhatsApp (whose role in spreading disinformation cannot be ignored).

- Content shared on one or two platforms = 0 points
- Content shared on more than two platforms = 1 point

## 4. Diffusion across communities: language as a proxy indicator

Inspired by Ben Nimmo's concept of communities used in his breakout scale, the language(s) in which the hoax appears will be set as an indicator of the spread of content in more than one big collective or group. The single piece of disinformation will score one extra point if it circulated in more than one language. In this regard, the researcher's lack of knowledge of foreign languages can potentially be an obstacle, but we believe that collaborative efforts between researchers and networks of fact-checkers can offer a way around the issue.

- Content circulated in one language = 0 points
- Content circulated in more than one language = 1 point

## 5. Media outreach: reaching mainstream media

The Breakout Scale also raises the question whether the content reached mainstream media. The researcher will score one extra point if the content has reached any mainstream media (e.g., quality newspapers, TV, radio) or if these media are the primary transmitters of the hoax, given their great outreach and expectations to uphold fact-checking standards. We will exclude junk sites, alternative media, or fringe sites that pretend to be authentic news media, but we will include instead state-sponsored media.

- Content did not reach mainstream media = 0 points
- Content reached at least one mainstream media = 1 point

## 6. Type of actor: public figures and persistent disinformative transmitters

The researcher will give one extra point if the transmitter or amplifier of the disinformative content is an influencer, a politician, an artist, or any public figure. Also, the hoax will receive an additional point if the transmitter has been fact-checked more than once, making them a 'repeated offender'.

- The transmitter/amplifier is not a public figure of any sorts = 0 points
- The transmitter/amplifier is a public figure and/or a recurrent disinformant who has been fact-checked before = 1 point

## 7. One narrative, many styles: appearance in various formats

The different formats in which a hoax is disseminated make it possible to reach further audiences. For this reason, we will give an extra point if the same hoax circulated in different formats, e.g., video, pictures, or memes.

- Content spread in one format exclusively = 0 points
- Content spread in more than one format = 1 point

## 8. Call for action and danger of the narrative: a multiplier effect

The fact that the hoax contains a call to action (i.e., and exhortation to organise a real-live protest or to harass somebody) could be a game-changer for the impact risk. Due to its importance, this indicator will have a multiplier effect on the first measurements regarding engagement on social networks. If the hoax includes an exhortation, the researcher will multiply the value of engagement by one. The result is then summed to the other values of the index. In this way, the risk that the combination of high engagement and an exhortation will reflect in the call-to-action measurement. If there is no exhortation, the call for action indicator will score 0. Therefore, the value of engagement is multiplied by 0, adding 0 points to the values of the index.

The content does not contain an exhortation = 0 points are attributed to the call to action (the value of engagement is multiplied by 0) and 0 points are summed to the scale.

- Engagement is 0. Then, engagement x exhortation is  $0 \times 0 = 0$  call to action: 0 points
- Engagement is 1. Then,  $1 \times 0 = 0$  -> call to action: 0 points
- Engagement is 2. Then,  $2 \times 0 = 0$  -> call to action: 0 points
- Engagement is 3. Then,  $3 \times 0 = 0$  -> call to action: 0 points

The content contains an exhortation = 0-3 points are attributed to the call to action (the value of engagement is multiplied by 1) and 0-3 are summed to the scale.

- Engagement is 0. Then, engagement x exhortation is  $0 \times 1 = 0$  call to action: 0 points
- Engagement is 1. Then,  $1 \times 1 = 1$  -> call to action: 1 point
- Engagement is 2. Then,  $2 \times 1 = 2$  -> call to action: 2 points
- Engagement is 3. Then,  $3 \times 1 = 3$  -> call to action: 3 points

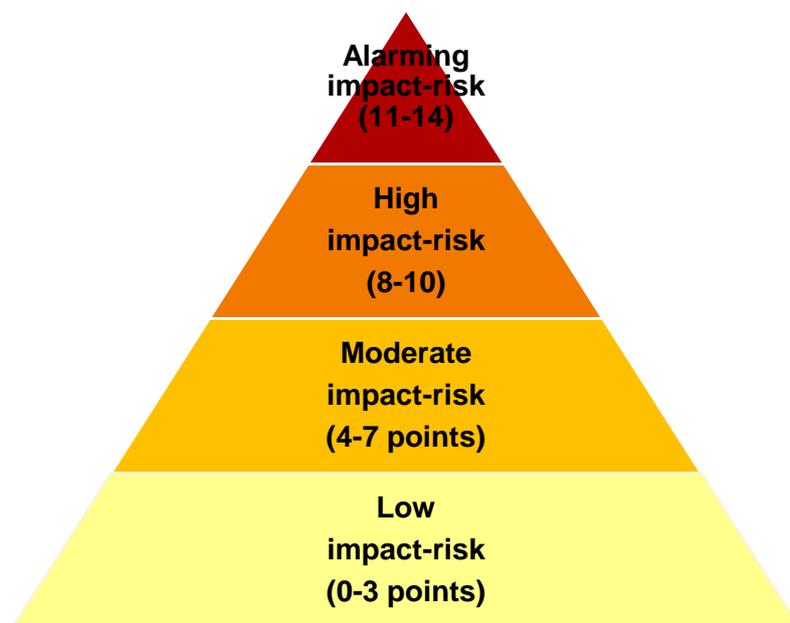
# THE FINAL SCORE AND THE SCALE

Applying this points-system, the hoaxes can score 11 points maximum without the multiplier effect, and a maximum of 14 with the multiplier effect of the call to action.

From 0 to 10 we will classify the hoaxes in low, moderate, and high risk while from 10 to 13 it will be an alarming impact risk, as you can see below.

| INDICATORS             | MEASURES  | POINTS                     |
|------------------------|---|----------------------------|
| 1. Engagement          | 0 – 1.000 shares and reactions = 0 points<br>1.001 – 10.000 shares and reactions = 1 point<br>10.001 – 100.000 shares and reactions = 2 points<br>More than 100.001 shares and reactions = 3 points   | 0-3                        |
| 2. Exposure            | 0 – 1.000 views = 0 points<br>1.001 – 10.000 views = 1 point<br>10.001 – 100.000 views = 2 points   | 0-2                        |
| 3. Number of platforms | Content shared on one or two platforms = 0 points<br>Content shared on more than two platforms = 1 point  | 0-1                        |
| 4. Number of languages | Content circulated in one language = 0 points<br>Content circulated in more than one language = 1 point   | 0-1                        |
| 5. Media outreach      | The content did not reach mainstream media = 0 points<br>The content reached at least one mainstream media = 1 point  | 0-1                        |
| 6. Type of actor       | The transmitter/amplifier is not a public figure of any sorts = 0 points<br>The transmitter/amplifier is a public figure and/or a recurrent disinformers who has been fact-checked before = 1 point   | 0-1                        |
| 7. Formats             | Content spread in one format exclusively = 0 points<br>Content spread in more than one format = 1 point   | 0-1                        |
| 8. Call to action      | The content does not contain any exhortation = 0 points<br>Engagement is 0. Then, engagement x exhortation is $0 \times 0 = 0$ call to action: 0 points<br>Engagement is 1. Then, $1 \times 0 = 0$ -> call to action: 0 points<br>Engagement is 2. Then, $2 \times 0 = 0$ -> call to action: 0 points<br>Engagement is 3. Then, $3 \times 0 = 0$ -> call to action: 0 points<br><br>The content contains an exhortation = 1 point with multiplier effect:<br>Engagement is 0. Then, engagement x exhortation is $0 \times 1 = 0$ call to action: 0 points<br>Engagement is 1. Then, $1 \times 1 = 1$ -> call to action: 1 point<br>Engagement is 2. Then, $2 \times 1 = 2$ -> call to action: 2 points<br>Engagement is 3. Then, $3 \times 1 = 3$ -> call to action: 3 points | 0-3 with multiplier effect |

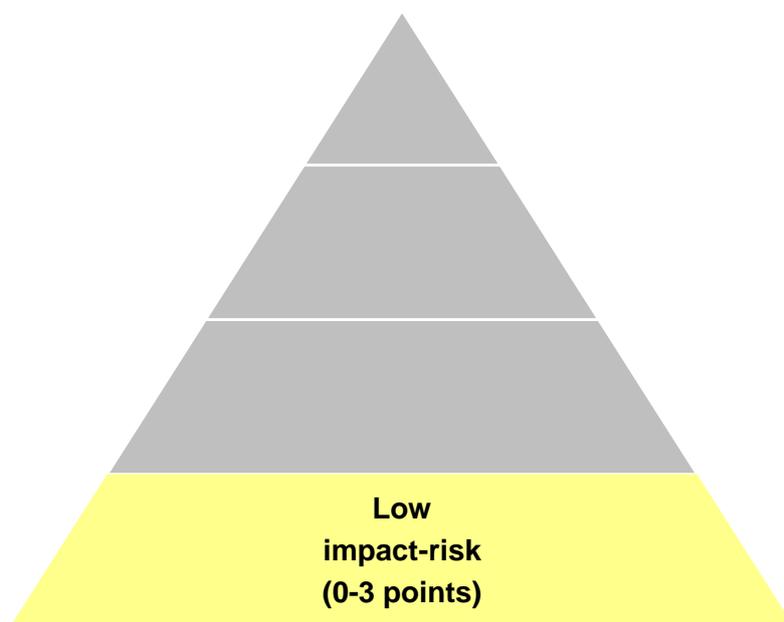
# INDEX APPLICATION OF FACT-CHECKED HOAXES



# Example 1

[Photo](#) showing sheep being driven to the slaughterhouse in three-store trucks was said to be taken in Germany, when it was actually from Australia in 2010.

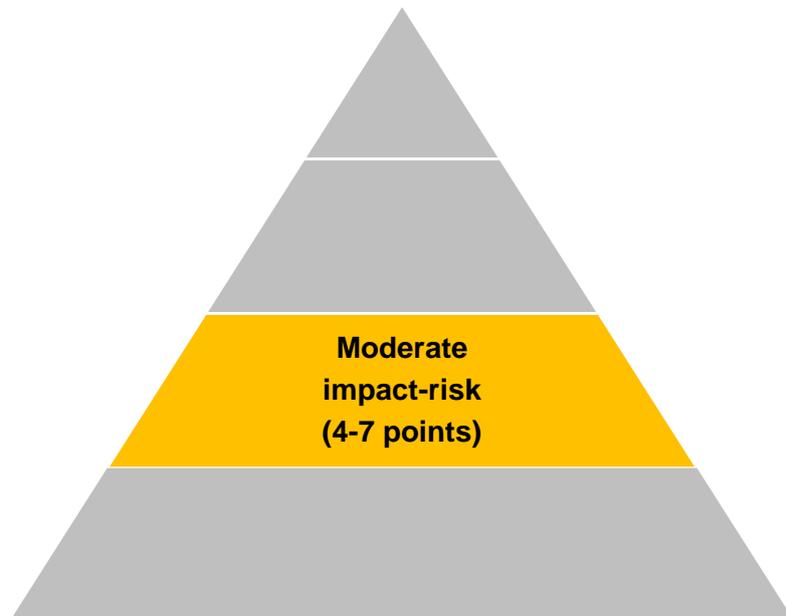
| INDICATORS             | MEASURES   | POINTS  |
|------------------------|--|---------|
| 1. Engagement          | 1.300 shares on Facebook                         | 1       |
| 2. Exposure            | Non applicable                                   | 0       |
| 3. Number of platforms | Facebook   | 0       |
| 4. Number of languages | German   | 0       |
| 5. Media outreach      | No   | 0       |
| 6. Type of actor       | No public figures or recurrent disinformers      | 0       |
| 7. Formats             | Photo  | 0       |
| 8. Call to action      | No exhortation (engagement x exhortation = 0 x0) | 0       |
| TOTAL                  |  | 1 point |



## Example 2

A fabricated claim was made that according to an article in Bild, [Ukrainian refugees](#) tried to burn a Russian flag in Germany, but ended up burning down their shelter instead.

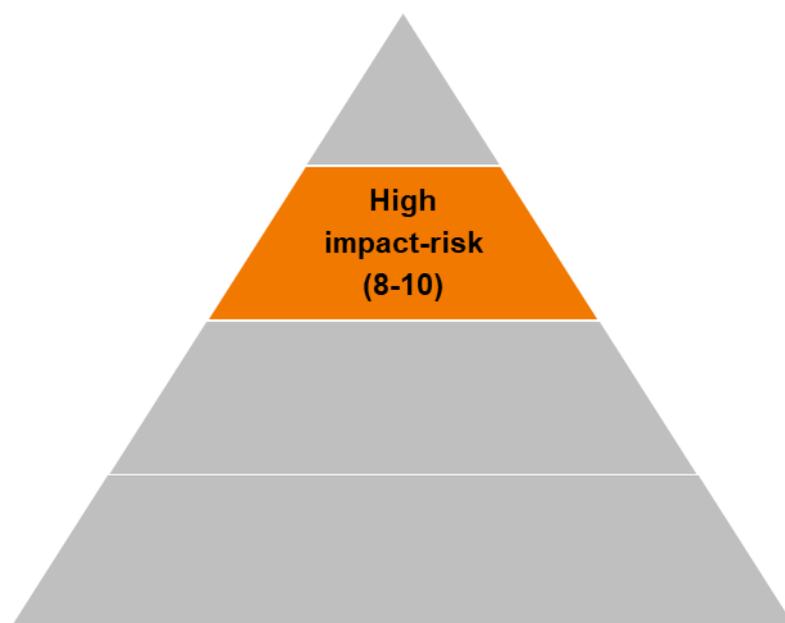
| INDICATORS             | MEASURES   | POINTS   |
|------------------------|--|----------|
| 1. Engagement          | 450 retweets                                     | 0        |
| 2. Exposure            | 67.000 views Telegram                            | 2        |
| 3. Number of platforms | Twitter, Facebook, YouTube, Instagram, TikTok    | 1        |
| 4. Number of languages | German, Russian, and Bulgarian                   | 1        |
| 5. Media outreach      | No   | 0        |
| 6. Type of actor       | Recurrent disinformers                           | 1        |
| 7. Formats             | Photo and video                                  | 1        |
| 8. Call to action      | No exhortation (engagement x exhortation = 0 x0) | 0        |
| TOTAL                  |  | 6 points |



## Example 3

World Health Organization leader, [Tedros Adhanom Ghebreyesus](#), has said that Covid-19 vaccines kill children and declared "Don't vaccinate your children!" In reality, he criticised some countries for boosting children, while in others, the vaccine is in short supply.

| INDICATORS             | MEASURES                                     | POINTS          |
|------------------------|--|-----------------|
| 1. Engagement          | 4.600 likes on Instagram                     | 1               |
| 2. Exposure            | 745.000 views on YouTube                     | 2               |
| 3. Number of platforms | Facebook, YouTube, Instagram, and Telegram   | 1               |
| 4. Number of languages | German, Spanish, and French                  | 1               |
| 5. Media outreach      | No   | 0               |
| 6. Type of actor       | Recurrent disinformers                       | 1               |
| 7. Formats             | Video and screenshot                         | 1               |
| 8. Call to action      | Exhortation (engagement x exhortation = 1x1) | 1               |
| <b>TOTAL</b>           |  | <b>8 points</b> |



## Example 4

Despite the lack of scientific evidence, a doctor declared that [Ivermectin and hydroxychloroquine](#) are valid treatments for Covid-19.

| INDICATORS             | MEASURES                                     | POINTS    |
|------------------------|--|-----------|
| 1. Engagement          | Millions of shares                           | 3         |
| 2. Exposure            | Millions of shares                           | 2         |
| 3. Number of platforms | Multiple platforms                           | 1         |
| 4. Number of languages | Multiple languages                           | 1         |
| 5. Media outreach      | Yes  | 1         |
| 6. Type of actor       | Public figures and recurrent disinformers    | 1         |
| 7. Formats             | Multiple formats                             | 1         |
| 8. Call to action      | Exhortation (engagement x exhortation = 3x1) | 3         |
| TOTAL                  |  | 13 points |

