

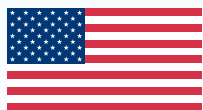


CIVIC  
RESILIENCE  
INITIATIVE








# DISINFORMATION TOOL-KIT

101 ON HOW NOT  
TO BE FOOLED





FIND MORE ABOUT **CRI**

-  [www.cri.it](http://www.cri.it)
-  CIVIC RESILIENCE INITIATIVE
-  @CivicResilience
-  @CivicResilienceInitiative
-  @civicresilienceinitiative

This publication was funded by a grant from the United States Department of State. The opinions, findings and conclusions stated herein are those of the authors and do not necessarily reflect those of the United States Department of State.



**IN THE PAST 10 YEARS, THE GROWTH OF SO-CALLED 'DIGITAL MEDIA' HAS BEEN EXPONENTIAL. WE ARE NOW FLOODED BY INFORMATION COMING FROM THOUSANDS OF DIFFERENT CHANNELS OF COMMUNICATION: MAINSTREAM MEDIA, ELECTRONIC PUBLICATIONS, WEBSITES, BLOGS, AND OF COURSE A WHOLE ARRAY OF SOCIAL MEDIA.**

As a result, we are free to choose the source which best reflects our interests, and political or social opinions. However, just as social media are increasingly becoming the primary source of information for many, they have also become a minefield through which disinformation travels rapidly.

With news-like formats being easy to reproduce, it is ever more difficult to spot the difference between a genuine article and a hoax. Genuine stories are twisted to push a political agenda, old video footage is used to illustrate recent events, edited photos are published to increase sensationalism of invented stories, armies of trolls and bots are deployed to attack an organisation or a particular public figure; it is becoming increasingly difficult to distinguish which news is real and which is fake. In order to fight disinformation, Civic Resilience Initiative (CRI) based in Vilnius (Lithuania) working with the U.S. Mission to NATO decided to produce the first 'CRI Disinformation Tool-Kit', first of its kind in the Baltic States. The aim of this Tool-Kit is to help journalists, media representatives and the civil society with a handy document that helps building digital literacy. As no equivalent is currently available in the Baltics, we are hoping this initiative will spark structural changes in society's digital resilience field. The 'CRI Disinformation Tool-Kit' is developed in cooperation with the two most prominent organizations in the field of disinformation: Baltic Security Foundation (Latvia) and Education of Media, Communication and Critical Thinking (Estonia).

CRI team has raised a goal to be the main catalyst in Lithuania and the Baltic region to strengthen the digital resilience of the society. Together with a digital resilience and a broader understanding of how to stay safe and vigilant online, this Tool-Kit will provide the needed tools and knowledge for less experienced information users on how to check, verify and select trusted information online for their daily use. Making sense of this new environment is far from simple. The challenge is to identify 'news' that can be misleading and prevent its further dissemination.

## **BUT THERE ARE MANY FORMS OF FAKE NEWS:**

- **DISINFORMATION:** Information that is false and deliberately created to harm a person, social group, organisation or country.
- **MISINFORMATION:** Information that is false, but not created with the intention of causing harm.
- **MAL-INFORMATION:** Information that is based on reality, used to inflict harm on a person, organisation or country.

All three types of information are dangerous because they travel far and fast, because they go 'viral': this happens when many people, and even organisations, repost these stories because they seem interesting and sensational, without giving them much thought. While we cannot prevent these stories from being created, we can learn how to detect a true story from a false one. By being more digitally aware, we can learn how to navigate in the confusing universe of millions of news stories, distinguish true stories from fake ones and make sure we do not inadvertently contribute to further disseminating them.

**WE ALL NEED TO BECOME MORE CAREFUL READERS OF THE NEWS WE ARE EXPOSED TO AND THIS CAN BE DONE BY LOOKING OUT FOR A FEW SIMPLE SIGNS. THIS BOOKLET OFFERS A NUMBER OF BASIC CHECKS THAT WILL ALLOW YOU TO FILTER AUTHENTIC NEWS REPORTS FROM THE ONES THAT DISTORT THE TRUTH. SPECIFICALLY, IT WILL GIVE YOU TOOLS TO:**

- **Check if online information is real or fake**
- **Identify trolls**
- **Identify bots**
- **Identify fake social media accounts**
- **Spot doctored images online**
- **Spot manipulated videos**
- **Guide you on what to do when you spot disinformation**

We hope that this short document will go some way towards increasing our resilience to the spread of malicious, fake information in the digital age, honing our skills in identifying and removing incorrect information and supporting the dissemination of reliable sources.



# IDENTI- FICATION

**How to verify  
news or posts?**



# How to verify news or posts?

If you read news and especially post on social media, where news or articles have been shared, it may be difficult to decide if the content is true or not. In cases where the topic is scandalous or sounds incredible, you should take a minute and do a simple check to verify if the information is correct and trustworthy.

## How to:

Here's five simple steps, which pieces of information you should check:

### 1. RATE THE SOURCE

Explore the website or social media account. Think about who might be behind the distribution of the news and what was the purpose of the story.

### 2. READ PAST THE HEADLINE

The headlines of stories can be scandalous, to attract clicks and promote sharing. If you dwell into the story, it may turn out that the claims in the headline are not true.

### 3. CHECK OUT THE AUTHOR

Does the named author really exist? Is the author a reliable person?

### 4. DO THE SOURCES CONFIRM THE STORY?

Often there are no links in the fake news that can be used to verify the facts. If there are references to the sources in the story, then click through them. It may become apparent that the original message has been embellished or the meaning distorted.

### 5. CHECK THE DATE

Re-publishing old news does not mean that they are still relevant.

## INFORMATION (TEXT) CHECKING

### How to check if text information is genuine?

Google, google, google!

To put it simply, a search engine such as Google is your best friend when it comes to disinformation. All of the information can be checked if it is true or false. The logic behind it to use the keywords you can identify from the information you are interested in and see if any reliable sources are already talking about it. If you are seeing something suspicious, it is a good chance someone is already talking about it.

A great advantage is that you can verify the information that is not digital as well in successfully identifying the keywords.

It is important to note, that it will allow you to find a reliable source with the same content you are trying to check, but you have to trust the source.

#### **How to:**

1. Identify keywords best describing the piece of information you are looking into;
2. Use one of the available search engines to look for the same information;
3. Identify reliable sources and verify the information;

#### **Mains rules to remember:**

- Googling for information is the best first step to check information. It is extremely fast and efficient;
- No matter what information you are interested in checking; Google text search can work very well: text, photos or videos;
- The key is to use keywords in order to find the same information on a reliable source;
- To be more efficient in googling, use the basic Google dorking operators, i.e. if you search a phrase, put it between quotation marks ("..."), if your search results has a very popular keyword, use minus (-) and this keyword to exclude it from the search results; if you don't know the accurate spelling of a word, or an exact number or a date, you can use \* as a wildcard to replace a missing character in a word or the entire word or a number or a date.
- If you speak another language, you can make good use of that. Search what other sources in other languages report on the issue.
- If you feel strong in one language, ask someone you trust to help you verify important information in other languages. Discuss with them how it is reported in other linguistic spaces. It may be as valuable for them as it is – professionally – for you.

#### **Useful tools:**

Google – [www.Google.com](http://www.Google.com)

Bing – [www.Bing.com](http://www.Bing.com)

Yandex – [www.Yandex.com](http://www.Yandex.com)

# IMAGE VERIFICATION

**How to check if visual information is genuine?**



## Reverse Image Search:

When it comes to fake images, the best methodology is to reverse image search. The idea behind the methodology is to use the search engine, but instead of keywords – use an image. This allows us to find all identical, or highly similar images posted before. As image recycling (posting an image from before and claiming it was taken recently) remains one of the main problems in the disinformation space, checking if the image was not posted before is one of the best strategies to counter it. Finding that the image was posted before is a reliable way to confirm that it was already on the internet. In other cases, if the image you are working with was altered, reverse image search can help you find the original image.

### How to:

- 1. Open one of the search engines (link provided next to the useful tools);**
- 2. Copy in the link or the downloaded image itself;**
- 3. Investigate if the same, or very similar images were posted before.**



# Error Level Analysis

Error level analysis (ELA) is a more advanced method that permits identifying areas within an image that are at different compression levels. With JPEG images, the entire picture should be at roughly the same level. If a section of the image is at a significantly different error level, then it likely indicates a digital modification. In practice, you should look around the picture and identify the different high-contrast edges, low-contrast edges, surfaces, and textures. Compare those areas with the ELA results. If there are significant differences, then it identifies suspicious areas that may have been digitally altered.

It is important to note, that this method is not bulletproof, nonetheless it is a reliable first step identification of digital alterations that were made to the image. Photo forensics is a separate scientific branch and to debunk well photoshopped images requires long years of skills, yet in terms of everyday propagandic messages, usually those images are not well designed and easily identifiable.

## How to:

1. Open Forensically or FotoForensics (links are provided below);
2. Upload the JPEG image of your interest;
3. Select ELA Analysis;
4. Look for inconsistencies in the analysis.

## Mains rules to remember:

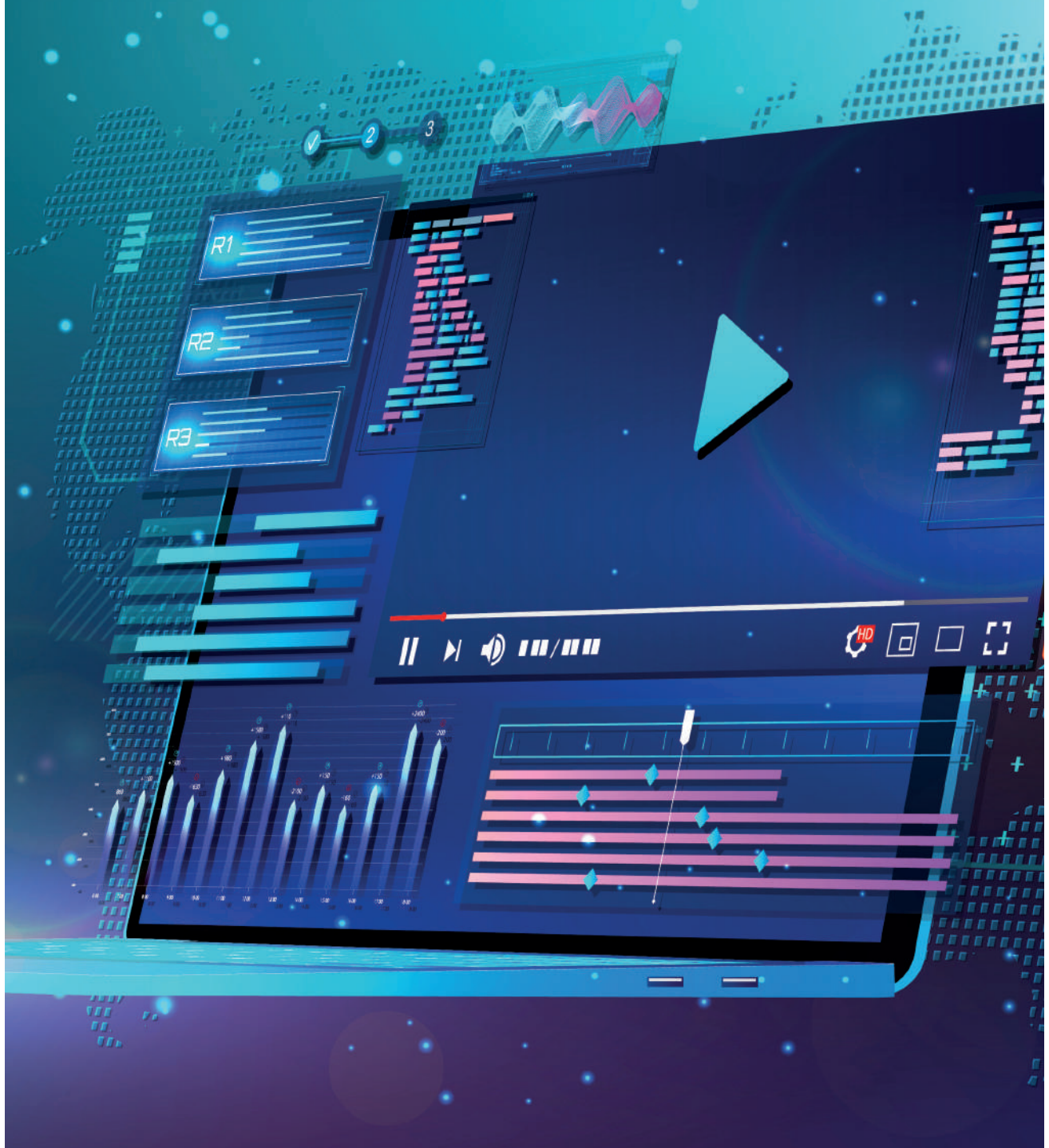
- Reverse image search should be a standard practice before trusting an image. If the image will not be what it said it was, it could deflect the actual genuine message;
- Reverse image search can be used to identify unknown people in the image;
- ELA is not a bullet-proof method, but it is a great quick way to check if the image is photoshopped.

## Useful tools:

Reverse Image Google Search - <https://images.google.com/>  
Reverse Image Yandex Search - <https://yandex.com/images/>  
RevEye Google Extension - <http://tiny.cc/ejcgdz>  
Forensically – <https://29a.ch>  
Foto Forensics – <https://fotoforensics.com>

# VIDEO VERIFICATION

How to check if the video  
is genuine?



# Reverse Image Search:

When it comes to fake videos, very similarly to images, the best methodology is to reverse image search. As videos are just a series of images, taking out a frame and searching for it is a great way to do it. Both tools InVid and Amnesty Data Viewer will allow you to find similar or identical videos already posted online, by looking for both frames and thumbnails.

## How to:

1. Open one of the search engines (Amnesty DataViewer or InVid);
2. Insert the link of the video;
3. Check if the video appears amongst duplicates.

# Forensic Analysis:

InVid also has the option of forensic analysis, when the software identifies potentially altered frames. The potentially altered frames pop-up in the analysis window, next to the sign 'Forensic'. If frames are identified by InVid as potentially altered, the chances that the video is fake are high.

## Mains rules to remember:

- The main threat with videos is the same as with images - video recycling. Videos taken before are being re-posted and presented as a fake piece of news;
- These tools are not as effective as image verification tools, yet they can capture a lot of the fake videos;
- Both tools are providing slightly different results, therefore for increased chance of recognizing fake videos it is a good idea to try both tools.

## Useful tools:

InVid – <https://www.invid-project.eu/>

Amnesty International YouTube Viewer – <https://citizen-evidence.amnestyusa.org/>

# BOTS

## How to spot a bot online?

### What is a bot?

"Bots" — automated social media accounts which pose as real people — have a huge presence on platforms such as Twitter. They number in the millions and are constantly being taken down and created again. These bots can seriously distort debate, especially when they work together.

Most typical uses of bots are:

- **to used to make a phrase or hashtag trend;**
- **they can be used to amplify a message;**
- **they can be used to attack a message;**
- **they can be used to harass other users.**

At the same time, many bots and botnets are relatively easy to spot by eyeball, without access to specialized software or commercial analytical tools. Software tools such as Sysomos, FireEye or Crimson Hexagon can be used to ease the search on a broader scale, but if you are in need to check whether the account is fake, **you can follow these 12 easy steps:**

**1. ACTIVITY** – is the most obvious indicator. This can easily be calculated by looking at its profile page and dividing the number of posts by the number of days it has been active. More than 50 posts a day can be regarded as suspicious.

**2. ANONIMITY** - the less personal information it gives, the more likely it is to be a bot. If there is no description, no image, or graph in the background – it is more likely to be a bot.





**3. AMPLIFICATION** - one main role of bots is to boost the signal from other users by retweeting, liking or quoting them. The timeline of a typical bot will therefore consist of a procession of retweets and word-for-word quotes of news headlines, with few or no original posts.

**4. LOW POSTS/HIGH RESULTS** - the bots above achieve their effect by the massive amplification of content by a single account. Another way to achieve the same effect is to create a large number of accounts which retweet the same post once each: a botnet.

**5. COMMON CONTENT** - the probability that accounts belong to a single network can be confirmed by looking at their posts. If they all post the same content, or type of content, at the same time, they are probably programmed to do so.

**6. SILHOUETTES** - The most primitive bots are especially easy to identify, because their creators have not bothered to upload an avatar image to them.

**7. UNORIGINAL PHOTO** - Other bot makers are more meticulous, and try to mask their anonymity by taking photos from other sources.

**8. ODD NAME** - A further indicator of probable botness is the handle that it uses. Many bots have handles which are simply alphanumeric scrambles generated by an algorithm.

**9. MULTILINGUALISM** - Some bots are commercial and seem hired out to the highest bidder regardless of the content. Such botnets are often marked by extreme diversity of language use.

**10. COMMERCIAL CONTENT** - Some botnets appear to exist primarily for that purpose, only occasionally venturing into politics. When they do, their focus on advertising often betrays them.

**11. AUTOMATION SOFTWARE** - Another clue to potential automation is the use of URL shorteners. These are primarily used to track traffic on a particular link, but the frequency with which they are used can be an indicator of automation.

**12. RETWEET AND LIKE RATIO** - A final indicator that a botnet is at work can be gathered by comparing the retweets and likes of a particular post. Some bots are programmed to both retweet and like the same tweet; in such cases, the number of retweets and likes will be almost identical.



#### **Main rules to remember:**

- The more criteria the account fits, the more likely it is a bot. Nonetheless if it fits at least two or three, it is enough to have an educated suspicion;

Be careful with calling accounts as bots in public. The better way to address them is bot-like accounts. This prevents false accusations as well as bot owners from tricking you by turning off the account automatization after identification;

- Software can be used to identify more strategic level trends, but the identification of individual

- Most of the bots are on Twitter, but they can appear on Facebook as well. Yet, most fake accounts on Facebook are more similar to trolls and will be discussed more in the "Trolls" section.

# TROLLS

## How to spot a troll online?

### What is a troll?

A troll is a person who intentionally initiates online conflict or offends other users to distract and sow divisions by posting inflammatory or off-topic posts in an online community or a social network. Their goal is to provoke others into an emotional response and derail discussions. A troll is different from a bot because a troll is a real user, whereas bots are automated. The two types of accounts are mutually exclusive.

To spot a troll is harder than to spot a bot, as these accounts are usually more sophisticated and are actively pretending to be real people. Below you can find a number of criteria that will help you to identify a troll, but these clues are indicative, rather than conclusive. It is seldom possible to say with 100 percent certainty that a given account belongs to a troll operation, rather than merely supporting certain malign narratives. Before examining the factors which reliably indicate a pro-Kremlin troll, it is important to look at one factor which does not - hyper-partisan content. A variety of contemporary real social media users tend to be highly partisan, especially when it comes to political topics.

#### 1. MISTAKES IN ARTICLES: A VS THE

One of the linguistic signs which is characteristic of many known Russian accounts is the inability to use the grammatical articles — “a” and “the” — appropriately. The Russian language has neither.

#### 2. MISTAKES IN FORMULATING A QUESTION

Another common linguistic indicator is the inability to phrase a question. In Russian, the word order for questions does not change, unlike in English, German, and formal French. Many known Russian troll accounts have posted questions which kept the word order of statements.

### 3. UNCLEAR OR QUESTIONABLE IDENTITY

Some trolls are using fake names that are very common in a given language, making it difficult to differentiate the specific author or deliberately leading to mistake it with another, such as a recognized journalist. The names used by trolls are also intended to be perceived as traditional or “sound right”, so that any reader would be tended trust or not to question such author of an article or a comment on social media. It may also be useful to check their profile pictures if there are any. Such pictures, added for achieving additional trust, may be stock photos that you may easily find on the internet. Such pictures may also be deliberately unclear when looking closer (photo editing, supposed person wearing sunglasses etc.), making it impossible to clearly identify the person.

### 4. AMPLIFICATION OF PRO-KREMLIN NARRATIVES

The Russian government has developed a distinctive narrative on key geopolitical events of the last five years. This follows the principles established as early as in the Doctrine of Information Security of the Russian Federation (2000) on conveying state policy and official position on issues that are important to the Russian government. As pro-Kremlin narratives are widely available on online sources, such as the Russian Ministry of Foreign Affairs or RT Twitter account, it is easy to check if the same themes appear in the suspected account. An account which repeatedly shares Russian government talking points on most or all of these events can justifiably be considered pro-Kremlin.

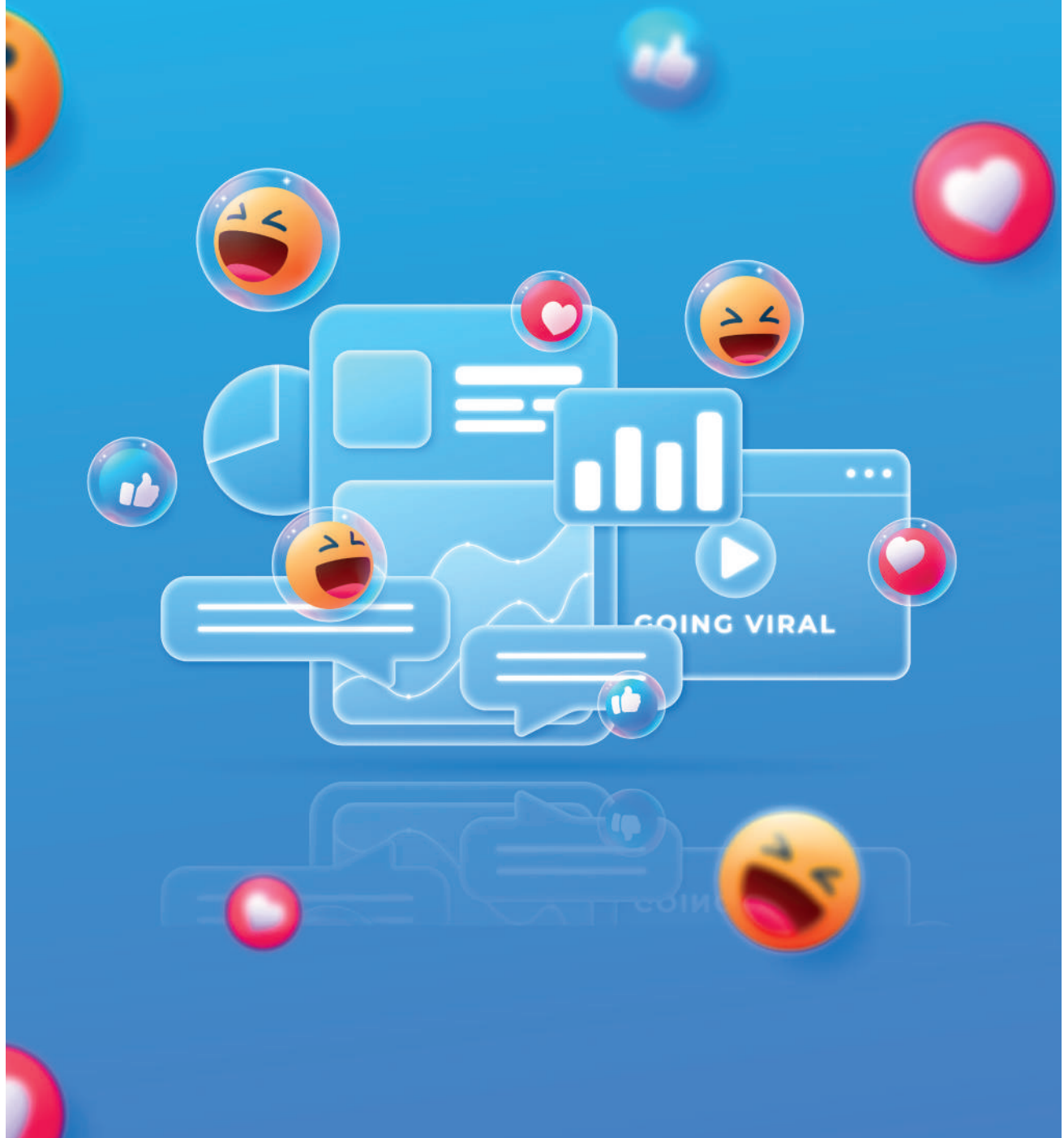
If the account shares most or all of the Kremlin narratives, makes the characteristic linguistic errors and poses as an American or British user, it may be a Russian-operated troll.

## Other Potential Clues:

- Trolls are anonymous - Most trolls use a nondescript first name, one that could be anybody — that is, if they care to use a name at all;
- Trolls have throwaway email addresses - As most places that allow comments require an email address, trolls get around this request by using made-up emails. Most of the throwaway emails are randomized and easy to spot, as they do not represent the person’s real name;
- Trolls are there to get a rise out of people - They’re not polite and not ashamed of getting in an open fight. They call names and make accusations and rarely do they sound anything but angry;
- Trolls use anonymous proxies - Trolls often use anonymizers, or proxies, that show a different internet protocol (IP) address than that which you’re used to;
- Trolls rarely add anything of value to the conversation - When trolls respond to a community discussion, they don’t add anything meaningful to the discussion. Instead, they joke, berate, and insult.

# FAKE FACEBOOK ACCOUNTS

How to spot if the Facebook account is fake?





Another important thing to have in mind are fake accounts. Usually, these accounts are not as active as trolls and tend to be more of a silent spectator. Similar criteria apply to fake accounts on most platforms, but in this guide Facebook was chosen as the main example. Facebook seems to be the most important, as users tend to share the most personal information on these accounts. These accounts actively try to become your friends for two main reasons: to appear more real, by having a number of real people as friends and to be in the friends list to see more personal information. Depending on the goals of the fake account, it can be used to collect personal information of the organization's employees.

#### **1. The factor of attractiveness.**

Not all unknown attractive young men or women social media accounts that invite you as friends are fake, but the vast majority of fake accounts seem to be. It is a higher chance that they will be accepted as friends and end up seeing your personal data.

#### **2. Few photo uploads.**

Most fake accounts don't post a lot of photos – three or four are typical, and occasionally they are pictures of different people. Just enough to create the temporary illusion that a real person is behind the account.

#### **3. Strange biographies.**

Most of the fake accounts have very scarce information in their biographies, or the provided information appears to be strange. For example, it is not impossible, but highly unlikely that a person is from the Bronx and attended the University of Helsinki, but also is very young and works for a New York PR firm. A quick check of their name on Google together with reverse image search of their profile picture can help you quickly debunk the fake account.

#### **4. Unresponsiveness.**

If you would reach out to a fake account, it is highly unlikely that it will answer even a short question. Ideally, it is better not even to try reaching out.

#### **5. A mostly blank Facebook Wall.**

Generally, the only things you'll find on one of these fake Facebook walls are new "Likes" on a Facebook company or product page and new friends.

# Response

To actively deal with online disinformation two parts are needed – to expose it and to report it. The organization should not bother to publicly acknowledge every fake story targeted at them and instead focus on positive and open messaging. A strong strategic communication explaining the events based on facts should be the strategic goal, rather than exposing and debunking the disinformation.

## **Making others aware;**

It is highly important to make your colleagues aware of the arising disinformation that targets the organization. Every organization should have a clear process in place for the employees to know where to send a report of a story. The main goal of this step is to make your colleagues aware that a certain circulating message is false and prevent them from sharing and believing the story.

## **Facebook/Twitter/Media outlet report;**

The second step is to report it on the social media platform. All of the social media platforms have the option of reporting a story with a specific reason why it is being reported. If the social platform receives enough flags from users, the story or the message will be taken down. This is the method that the grass-root civic organizations (a.k.a. Elves) are using to fight disinformation online. If media outlets are pushing fake stories, depending on the nature of the media outlets (if they are genuine or propaganda outlets), it should be reported either to them or to national media control institutions.

## **Advanced tools and reaching out for help;**

When it comes to harder disinformation cases, two main approaches are available: use more sophisticated open source methods or reach out for help to the online research community.

Most of the online tool box are relatively easy to use and provide step-by-step instructions on how to use them. Two large and the most useful tool boxes are provided below:

**Bellingcat Tool Box:** <http://bit.ly/2nBYzF>

**Online Open Source Tool Box:** <https://start.me/p/Wrrzk0/tools>

The other option is to reach out to the online research community and provide them with leads on a certain story. Most of the researchers will be happy to debunk the story and to share it online.





---

Vilnius, Lithuania  
Edition – 1000  
©CRI, 2021