

Social Media Use and Fact - Checking

Contents

Interacting Online Mis-, Dis-, and Mal-Information

- Information Verification Tools
- Spotting Mis- and Disinformation
- Verifying Images
- Africa Fact-Checking Fellowship #AFFCameroon

Online Conduct

- Social Media Code of Conduct
- ThinkB4UClick
- Engaging with Hateful Messages
- Additional Resources



*Africa Fact-Checking Fellowship fellows in session.
Photo by Derick Kinang*

Interacting Online

In the last chapter, we talked about hate speech and its harmful effects.

The next question is: **how do we work towards countering hate speech and spreading messages of peace and reconciliation?** That's what this chapter will focus on.

We will discuss what misinformation and disinformation are and how to spot them. And we will talk about guidelines on how to engage with the online world in a positive, peaceful manner.

We all hold the power to stand against hate and be responsible stewards of our online life. Let's learn how to do this well.

Mis-, Dis-, and Mal-Information

Online, it is very easy to share and repost other people's content. That ease of posting can be a good thing, but it also has negative consequences: a lot of false information is posted and shared.

But not all false information is posted with the same aim. There are different levels of falsehoods, called misinformation, disinformation, and mal-information. What is the difference? In essence, the difference is the intent behind why the wrong information is being shared.

Misinformation is when you share incorrect information, but you do not know that it is wrong. This means that the person posting misinformation is not intending to cause harm. They are simply ill-informed. But even if misinformation is not done with bad intent, it still causes harm. The good thing is because misinformation is based on people having wrong facts, it can easily be countered and corrected.

An example: You share information about electricity prices doubling soon because you heard this from a close friend. You failed to verify this information and it is not true.

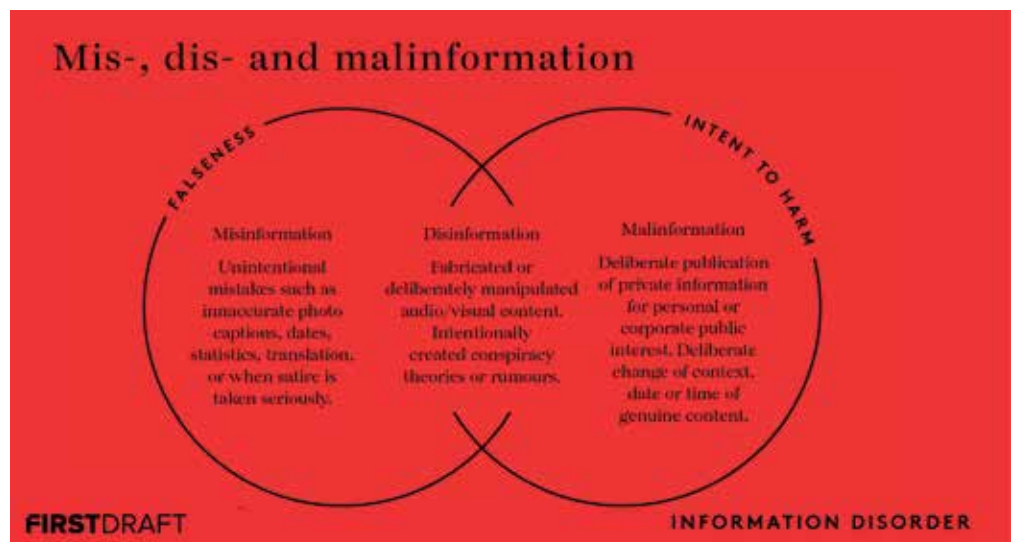
Disinformation is when you purposefully share incorrect information in order to deceive people. This means that the person posting disinformation is intending to mislead, deceive, and cause harm. Disinformation cannot be fixed by simply presenting facts, because the person

already knows it is wrong. A lot of disinformation is often made up of a mix of truth and lies, where lies are mixed in with truths. This is done so that the person reading the disinformation feels that since some of the information is familiar or true, the rest must be as well. Beware of this trap.

An example: You see a post that a politician supports closing a local school, which is a lie. The person sharing this knows that this is not true but wants you to vote against the politician and so is spreading disinformation to influence your vote.

Mal-information is information that might be true or not but is shared solely with the aim to inflict harm, either towards an individual or a larger group.

An example: Someone shares a photo of a dead child and posts saying that members of another tribe killed the child. The photo might be real, but it could be from a different country and it is shared solely to incite anger against the other tribe.



Why is it important to spot mis-, dis- and mal-information? Because even small falsehoods hold the power to hurt people and lead to conflict. The more we learn to demand the truth, the safer our world becomes.

The problem is that we are not trained to spot mis- and disinformation easily. This is a skill we need to learn how to do. We will cover this in the next section.

► QUESTIONS TO CONSIDER ◀

- Can you, in your own words, describe the difference between misinformation and disinformation?
- Can you give an example of how false information can cause harm to you or your community?
- Who benefits from false content?

Information Verification Tools

One of the objectives of the #defyhatenow initiative is to only spread truths online. But in a world where it is so easy to share and repost content, **how do we know if what we see is true?**

The answer is that it takes effort and rethinking how we interact with the world. We have to question everything we see online, even if the information comes from people we know and trust.

Keep in mind that misinformation is wrong information spread by accident - the person doesn't know it is false information. So that means that even when we hear something from someone close to us, we have to use these same tools that start with the motto **Question Everything**.

Spotting Mis- and Disinformation

Below are a few steps that you can take to try to spot mis- and disinformation when you see something posted as fact online.

Research the source: Who's sharing this information? If it's online, does the website sound familiar? Does it have political affiliations? If it's a person sharing, is this individual a frequent exaggerator? Do they have a job, education, or experience that would give them insider info? Do they seem trustworthy? What do you know about them?

Question the motivation: If you had to guess, why do

you think this person is sharing this content? What is their motivation? Is the story balanced and fair? Who gains from you reading the post and acting on it?

Check the date: It is easy to get angry about an article online that someone shares, only to realize it was published years ago and no longer applies. Along with the date, check the location and other facts to see if it is actually relevant to your situation.

Cross-check: If someone is sharing "news", check if other reliable news sources are reporting the same information. If not, it's unlikely to be true. Also, keep in mind that there are more and less reliable sources. When it comes to news, websites such as Al Jazeera, AllAfrica, and The BBC are very trustworthy. If you are looking for facts, a good site is Wikipedia. Wikipedia is a website that collects information on many topics and is peer-reviewed, meaning that many people work on the information and have all agreed that it is correct.

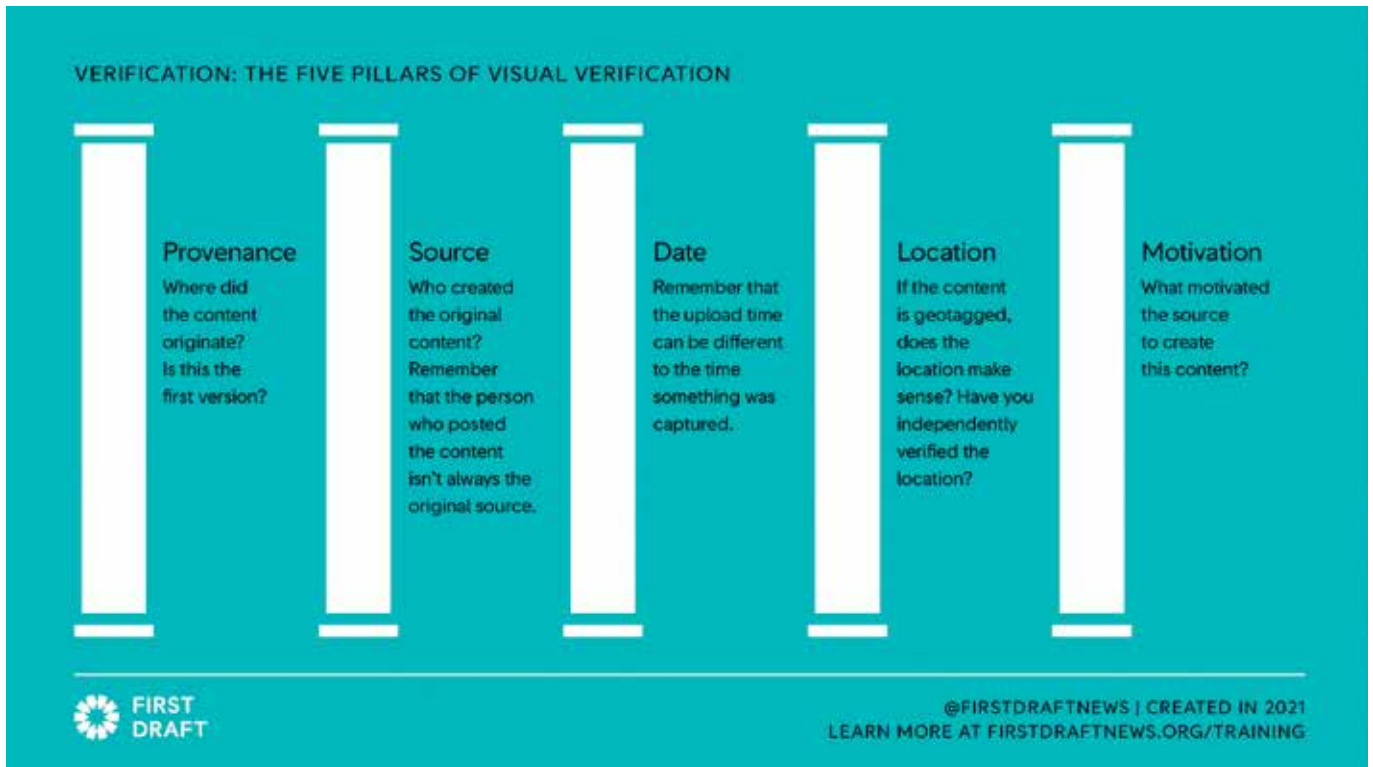
Read past the headline: People create headlines that sound more intense and scandalous than the story really is. Read past the headline and see what the entire article says. When reading, see if the article is listing its sources. And if they are, verify that those sources are accurate. If there is a story and no information on the source, there is a good chance the story is not true.

Question emotionally charged content: Does the post contain emotionally manipulative language designed to get you upset or excited? That's a red flag. Reliable sources let the facts fuel your response, not emotional language.

Distrust absolute narratives: If you come across a post or story claiming something is only good or bad, be careful. If the content sounds too good to be true, it probably is. If the content sounds only bad, it probably is not accurate either. Look for a balanced representation of reality.

► QUESTIONS TO CONSIDER ◀

- When you see a post online, is your first reaction to question whether it is true?
- Which of the tools above feels easiest to apply? Which requires the most work?
- What are some red flags that tell you a story might be fake?
- How could you encourage others to verify content more closely?
- Could you teach your family members and friends about what you have learned and the steps they could take to fact-check information?



Verifying Images

Now that we have learned about how to spot falsehoods in the content of posts, let's also talk about how to verify images.

Images are difficult to deal with, because visuals can be manipulated and changed using computer programs, so what you see in a photo may or may not be real. Additionally, mis- and disinformation is spread by sharing photos out of context, for example by posting a photo from a war scene, but it shows a war from a different country or time period.

Below we offer some ideas of how to verify whether images you are seeing are fitting and truthful or whether they are likely trying to trick you. Regardless of whether you are looking at a photo from a close friend or a stranger, the questions to ask are the same. These questions are similar to the questions we discussed above regarding content and have the same motto of **Question Everything**.

Provenance: Are you looking at the original photo, taken by the person posting? Or has it been reposted many times and the original photographer is no longer known?

Source: Who is posting the photo now? How did they get the photo?

Date: When was the photo taken? Do you have any way of knowing this based only on the photo?

Location: Where was the photo taken? Look at it closely and see if it looks like the location it is pretending to be.

For example: Do the streets and houses look like they should? What side of the street are cars driving on? What do the trees and plants look like? How are the people dressed? Look for clues to see if it fits what it says it shows.

Motivation: Why is the person posting this photo? What is their motivation? Take a guess about how they are trying to make you feel and what reaction they are trying to get out of you by posting this photo.

Here is a more detailed question guide that can help you assess the risk that you are being misled by an image. It moves from danger (the word NO and red color) to trustworthy (the word YES and blue color).

Another way to verify images is to use Google Chrome **image search**. Just as you can use Google to check for facts and names, you can also use it to see whether an image has been previously published online. How? Right-click on any image, hit "search google for image," and see if a picture has appeared online before.

VISUAL VERIFICATION GUIDE: PHOTOS

	NO				YES
1. ARE YOU LOOKING AT THE ORIGINAL VERSION?	A reverse image search returns identical photos indexed online before the event in question took place	A reverse image search returns similar results with some identical features, suggesting it might be a composite	A date search on each social network reveals it to be the first of many versions shared online but we have not yet received confirmation from the uploader	We are unable to find other versions online and basic shadow and reflection checks suggest that it has not been manipulated	It was sent to us directly and we have spoken to the source
2. DO YOU KNOW WHO CAPTURED THE PHOTO?	It came in via an anonymous email or chat app number	It was uploaded to a social network but the username does not appear elsewhere online. The uploader wants to remain anonymous	By running full name searches, reverse searching the user's profile photo, and/or researching the domain ownership of their blog or website, we were able to identify the uploader	We communicated with the uploader via social media to confirm that they took the photo	We questioned the source and their answers correlated with EXIF data, weather reports and their own online footprint
3. DO YOU KNOW WHERE THE PHOTO WAS CAPTURED?	There was no location data available and it contains no visual clues to investigate	We have cross-referenced with other photos coming from the scene but there is no satellite or street view imagery available to confirm the location	We have used visual clues such as signage, architecture and clothing to establish a broad geographical region	We have crossreferenced landscape and landmarks using mapping tools and have confirmed the lat/long coordinates	The source was able to confirm other landmarks in their field of vision, which matched those shown on online maps
4. DO YOU KNOW WHEN THE PHOTO WAS CAPTURED?	It was sent to us anonymously and there is no EXIF data available	We checked the timestamp on the social network to see when it was first shared online but we have no EXIF data confirming when it was taken	The social time stamp shows it was uploaded shortly after the event occurred and it features visual evidence that correlates with other eyewitness reports	We confirmed that the weather conditions and any shadows visible in the image correlate with the time, date and location given by the source	It contains EXIF data that, combined with other checks, confirms when it was taken
5. DO YOU KNOW WHY THE PHOTO WAS CAPTURED?	We do not know who took the photo so we can't ascertain what their motivations might have been	The social media account was created very recently and/or social searches reveal the uploader rarely posts online so there is little evidence to confirm their movements or motivations	Wider online searches of the uploader's real name reveals that they are connected with an activist or advocacy organisation but there is no additional information to know their motivation in this case	Searches of the uploader's social media activity leading up to the event confirm their reasons for capturing the photo, i.e. holiday-maker, journalist, works locally	The photographer confirmed the circumstances surrounding the photo

► QUESTIONS TO CONSIDER ◀

- When you see a photo posted online, is your first reaction to wonder whether it is real or manipulated?
- Which pillar of verification discussed makes the most sense to you? Which seems complicated?
- What are some red flags that tell you a photo might be fake or taken out of context?

**Africa fact-checking fellowship
#AFFCameroon**

The Africa fact-checking fellowship (#AFFCameroon) is an initiative by #defyhatenow, in partnership with Data Cameroon, to promote fact-checking, data journalism, and digital rights among journalists, bloggers, and online content creators. The quarterly

program aims to provide fellows with the skill sets and tools needed to tackle the current mis-, dis-, and mal-information challenges in the countries they live and work in.

#AFFCameroon uses a blended learning model that utilizes webinars, practical lessons, peer learning, and hands-on field assignments. The fellows who have participated in the training program often include online content creators, media journalists, and communication experts.

#AFFCameroon also offers fact-checking workshops, during which you can sharpen your fact-checking skills. More information can be found online.

Learning how to spot mis- and disinformation online helps you be a responsible social media user, because it prevents you from engaging with or sharing false content or images.

Online Conduct

The next step is only creating content that fosters peace and is accurate and fair. Keep in mind that anything you write, link to, and mention online is public and can be shared with anyone. Distinctions are generally not made between your 'private' and 'professional' opinions. Statements, tagging, or actions like 'liking' comments on Facebook will be interpreted as an indication of your personal opinions. That is why it is so important to **#thinkb4uclick**.

Below is a code of conduct that offers guidelines on how to engage responsibly with social media, both personally and professionally.

Social Media Code of Conduct

BE TRANSPARENT

When you post, you should identify whether you are posting on your behalf or on behalf of an organization. The person seeing your post should know why you are posting the content you are sharing.

BE CLEAR

Post clear messages and take responsibility for the content you post, both when people react positively and less positively.

CHECK YOUR FACTS

Before posting, always make sure your facts are correct. If you are reposting information from others, verify your sources. If you can't be 100% positive that what you are sharing is accurate, do not post it.

BE FAIR AND RESPECTFUL

Never post malicious, misleading, or unfair content. Do not post content that is obscene, defamatory, threatening, or discriminatory. Do not post comments that you would not say directly to another person. Always consider how other people might react before you post.

BE POLITE

Stay polite in tone and respectful of people's opinions, especially when discussions become heated. Show proper consideration for other people's privacy.

AMPLIFY THE POSITIVE

Encourage the good and help build connectedness, engagement, and community.

GIVE CREDIT

If you post something that originally came from someone else (like a quote, an image, or an idea), say so in your post.

▶ QUESTIONS TO CONSIDER ◀

- Can you agree to follow these guidelines? Do you already?
- Which guidelines are harder for you to follow? What makes them harder?
- Do you feel most people in your life post more positive or negative content?
- How could you encourage others to follow these guidelines?

#Thinkb4uClick

#defyhatenow launched the **#ThinkB4UClick** campaign to encourage people to be responsible social media users, to learn to verify information before they share misinformation, and to support peace. You too can join this movement! Become a peace ambassador by agreeing to follow the guidelines below.

As a Peace Ambassador on social media, I will:

- think before I click
- always verify information before sharing or reposting anything
- analyze posts I write to ensure that what I am saying is accurate and positive
- be responsible for my actions

Engaging with Hateful Messages

There is disagreement about whether it makes sense to respond to hateful messages online. Most often, the better option is just to amplify the positive and use your voice to foster peace. But if you do feel the need to respond, here are some guidelines offered by the Media Diversity Training Institute on how to respond to hate speech.

DON'T BE ABUSIVE

Make sure your words and any content you share do not spread bigotry, prejudice, or hate.

DON'T SPREAD THEIR HATE

Giving attention to fringe individuals and their hatred can be counterproductive. You might give them the attention and publicity that they crave.

DON'T RESPOND TO OLD POSTS

Responding to old posts risks reviving a conversation and having the adverse effect of spreading hate speech to others who might not have seen it before.

THINK ABOUT YOUR OBJECTIVES

Why are you wanting to engage with the person expressing hateful speech? Are you seeking to lessen the consequences of a hateful post on a wider audience? Are you seeking to express support for a group or identity under attack? Thinking about your goals will help shape your response and the language you use.

TRY TO ENLIST INFLUENTIAL SUPPORTERS

Add their usernames to posts. Celebrities, politicians, civic leaders, and subject experts can help bring attention or add weight to your counter-narrative.

DON'T FEED THE TROLLS

Some people are open to discussion. But most people posting hate messages are not seeking engagement; they just want to incite anger. These sorts of people are called trolls. Don't interact with them, because they are motivated by engagement.

BUILD A KNOWLEDGE BANK

Develop a resource bank of counter-arguments, statistics, information, sources, and links to support your counter-narrative. Subject knowledge is important and those engaging in hateful speech may be armed with many misleading sources of information that they will use. Counter them with credible evidence from an independent and reputable source.

Additional Resources

First Draft is a trusted source and has many resources and videos that go into more detail on how to verify online information.

<https://firstdraftnews.org/long-form-article/verifying-online-information/>

They also offer resources, instructional videos, and training videos on other related topics. <https://firstdraftnews.org/long-form-article/first-drafts-essential-guide-to/>

To fact check information found online, doing a

Google search is a good idea.

www.google.com

AfricaCheck is also a site that will help you sort facts from fiction.

www.AfricaCheck.org

The League of Women Voters in the USA offers a guide on mis- and disinformation.

<https://www.lwv.org/blog/your-guide-mis-and-disinformation>.

Safe Sisters offers tips and tools to safely interact with the digital world, including teaching you how to create stronger passwords.

<https://safesisters.net/>