

# MEDIA LITERACY TOOLS IN COMBATING DISINFORMATION AND FAKE NEWS IN SOCIAL MEDIA

Maria Magdalena POPESCU<sup>1</sup>

**Abstract:** *If mainstream media meant one sender and inordinate receivers, nowadays, new media brought the chance and challenge to have numberless senders and receivers at the same time, in a network of information. Along with polyphonic chunks, inaccurate information penetrates the echo-chambers we create. The danger is that unless timely spotted, it sows disinformation and polarization. The solution resides in media literacy skills, in raising awareness over types of communication products meant for malicious use especially within social media where user generated content contributes to proliferating and spreading the content at incredible speed. The present paper is meant to present ways that help individuals avoid being subject to mal-intended behavioural and cognitive influence.*

**Key words:** *media literacy, fake news, disinformation, social media*

## 1. Introduction

The growth in the number of apps used for socializing and information along with technological development have given rise not only to positive but also to negative or malicious use of social media. The automatic character, the repetition of information, the permutation and the production of user generated content, the anonymity, the impersonal character, the distortion to a totally different representation of the initial content are all characteristics that allow some users to turn social media into a dangerous tool. The growth of malicious users, virus writers and phishing attacks have come along with all the technological progress (Robertson, 2010). What is more, the widespread phenomenon of social media malicious use is affecting the public sphere, from governments to individuals, to influence knowledge and perceptions and to generate behaviours, in a direct or indirect way, through global access of rapidly distributed media due to permeable, ubiquitous information. Social platforms like Twitter or Facebook that once enhanced democratic approaches and open exchanges of ideas are now just resorts for manipulation and tools for spreading disinformation. Social media along with mainstream media using algorithms, automation and the

---

<sup>1</sup> Carol I National Defense University, [popescu.maria@myunap.net](mailto:popescu.maria@myunap.net)

Internet are now disseminating narratives that equally influence and impact governments and people, posing the agents involved in communication on different levels of power status.

## **2. Social Media Background**

There is a plethora of definitions and approaches regarding what social media is and what it consists of, most commonly being referred to as a variety of forums, blogs, chat rooms, rating services and microblogs (Freberg, 2019) kept functional through online practices that use technology and allow the distribution of content, opinions, experiences, media or insights. However, social media is also about people that not only establish relationships but also build their own amplified content or generate a new one based on conversations, dialogues and relations, in a hub of information, within a network where key audiences can be identified and persuaded to create and distribute content, in their turn, through impression management and reputation tools. Moreover, research has shown that while social media provides the feeling of a great degree of control over a situation, it stimulates users' desire to get involved and give support to the community through social participation and personal requirement (Whitmore, Agarwal & Da Xu, 2014), sometimes creating social overloads (Stenger & Cotant, 2010) in places where users aggregate based on common interests, on a large demographic spectrum. The environments thus created and inhabited are the cradle for social group interaction, for education, entertainment and most often for information that has lately been infused with disinformation techniques. Users tend to combine various media in a network of information and communication, expanding the content, generating a convergent culture (Jenkins, 2006) through trans-media storytelling (Scolari, 2009), a sharing of mundane and sometimes (Wilson, 2004) unimportant but polarizing (Malinen, 2015) meaning on various platforms. Other times, people just create their own information sources through the "daily me" selection process of certain types of information, enclosing themselves in echo chambers and filter bubbles that enhance the above mentioned polarization in the content spread, especially in the one with a strategic goal, in an international politics environment.

Examples include here social media bots used to spread disinformation in French elections through the Macron Leaks campaign (Ferrara, 2017), the ones used to advance hyper-partisan information during the Brexit referendum (Bastos & Mercea, 2017), or to influence presidential conversation in the 2016 US elections (Bessi & Ferrara, 2016). Moreover, Facebook and Twitter had to explain, in 2018, their involvement with digitally enhanced foreign manipulation in collaboration with mainstream media, related to the US elections as well. That is why, nowadays differentiation should be made between malign and benign use of new media in the growing information warfare. COVID-19 pandemic spread is evidence that world wars can be battled without conventional ammunition and still invisible warriors can fight weapons of mass destruction. Disinformation and manipulation as psychological operations can only weaken people's confidence in state institutions that in turn weakens homeland and national security. Lack of media literacy has thus become a soft spot across many societies and a growing

need most of all. Lack of awareness in methods used for disinformation and its products does not allow people to crosscheck the information in a world of abundant unreliability. Conversely, better management of indicators about what fake news is and how it can be differentiated from the real one, how we can fight the invisible weapon of disinformation, propaganda and informational operations can only be achieved with mastering media literacy and media awareness skills.

### **3. Malign Use of Social Media – Towards Media Awareness**

To fight disinformation from the first steps and seize the difference between malign and benign social media content, media literacy skills need to develop or to consolidate towards the methods used for generating the already mentioned harmful and deceitful information. These represent the cornerstone which allows us to understand how malign content is produced, helping one to identify it and thus avoid being subject to disinformation in this way. The following are among dangerous uses of media, especially social media that Robinson (2018) draws attention to in its mission for a correct perception of information. If used, these alter individuals' perception of the real intended message:

- 1) Impersonation, originally meaning pretending to be another person, according to Cambridge Dictionary (2020) in social media refers to disinformation and social engineering attacks that allow the creation of events similar to reality. Spread online, Artificial Intelligence contributed to the development of deep fake content, a type of content very similar to reality that represents pictures or videos of popular figures (meant to generate credibility) and induce the mal-intended change of behaviour. Impersonation can be applied not only to individuals but also to media content and new media content. The differences from the real model are extremely thin and almost invisible (a change in a letter to a website or a domain name in URL, a particular physiognomy mark in the fake photos or videos representing popular people, etc.). The hostile actors are effective in covering their traces. Impersonation can also be manifested through the malicious use of media and social media when users copy credible sources and exploit public trust to promote disinformation. Deepfake, as this is called, represents realistic digital manipulation of sounds, video images, while the difference between truth and fake is blurred and the cognitive biases escalated. True-like videos presenting unreal events are developed with machine learning and artificial intelligence to produce deepfake.
- 2) Social engineering attacks, another way to produce malicious content, target a huge amount of information displayed in social media while the security systems advance. Social engineering approaches are materialized in posts, apps that require personal data, data that can break the own accounts, extract users' information through creating opportunities for logging into various apps in exchange for providing birthday, mother's name, the street where you grew up, the name of your pet, etc. Thus, security questions of personal accounts are hacked and then used for other purposes. Romantic fraud is another tactic in which military personnel might seem to be looking for romantic involvement, asking for friendship through Facebook or

Instagram accounts. In reality they are software agents created to extract personal and especially financial information about the targeted person. An alarming sign, if such a case is encountered, is that those accounts do not have a number of friends or followers, do not have posts, they manifest themselves like a newbie, like someone new in the social media environment.

- 3) Fake news is prolific due to information abundance and technological variety. It was first fabricated with the mobile print, in 1439, due to the lack of journalistic and ethical integrity. The fake news was then a story about sea monsters, despite writers and educated people's fight to combat the phenomenon (Voltaire was preaching against religious fake news, while in America propaganda used scalping Indians for bloody images and control). Fake news continued to expand in the press as well, with the penny (yellow) press or the tabloid press- *The Sun*- in the US in 1830, as an example. At the time, the newspaper published news on life on the moon, with biped weasels and unicorn goats. In the same context, the well-known phenomenon of 1938, mass disinformation on Wells' radio theatre, the War of the Worlds, a simulation of a live transmission presenting New Jersey and New York attacked by Mars' aliens was manifested. Nowadays when one speaks about fake news one triggers three terms, simultaneously, i.e. *dis-information*, *mis - information* and *mal-information*. Disinformation is deliberately created to be false and induce harm while the persons who disseminate it are aware of the fact; mis-information, on the other hand, is false for any other reason, either by omission, by ignorance or through any other factors that intervened in the process of delivering the message in a truncated manner; the persons who spread misinformation are convinced that what they disseminate is true, even though the information is false. Ultimately, mal-information speculates real facts and is meant to do harm on individuals, society or organizations. Mal-information is about things that once made public can harm people (news about someone's beliefs on a certain topic) Generally taken with the encompassing term of fake news, all these instruments
- 4) Disinformation is nowadays aided by the tools in social media, bots and trolls, which bring violence and are mainly used in elections and political relations, in state-sponsored campaigns (see Russia and EU) The bots and trolls are hostile actors that exploit human biases and social media ecosystem vulnerabilities. Trolls concentrate on users engaged in controversial discussions and aim at spreading fury and disagreement. Bots are automated, anonymous and become visible due to intense activity, conventional material but no interaction with the other users. Internet Research Agency is a Russian factory for trolls in the US elections, a real warning due to the large scale coordination. The bots and trolls amplify certain narratives, they manipulate the information environment and make certain visions or political events be more popular than they really are- In this case, they bring comfort to dissidents by developing social acceptance for the stories they promote. Also, what they do is destabilize the public discourse, undermine cohesion, fuel chaos, making it difficult to understand where the truth is and what the fake is. The bots promote certain specific interests, augment the posts that show interest and gain influence in social media. A hostile use of social media that negatively affects through bots and

trolls targets the military. The bots are software agents that communicate autonomously in social media to influence discourse or users' opinions (chatbots) through simple and repetitive tasks. Similarly, the trolls create hate through expressing controversies and become visible in the crowd. Online agitators, they inflame, offend and attack sentimentally, not intellectually, engage emotions and determine the targets to take action.

- 5) Nowadays technological progress and information abundance make it difficult to differentiate fact from fiction. One needs selection abilities and critical evaluation skills. In the post truth era facts are less influent in shaping public opinion than the appeal to emotions and to personal belief. Moreover, with technological devices being more and more apprehensible, social media has given rise to citizen journalism, a collection, circulation, analysis and information, made by citizens. The collectable sets of events are then shared publicly without any confrontation with reality, without any cross-checking of facts versus opinions, giving easy rise to fake news and disinformation. Narrated in the first person as witness – the subjective news primary source is shared more often and on a larger scale in social media than the real news, due to emotional elements contained in the event, in a competition for exclusivity and desire to be the first. The core of disinformation is falsification of events, people, facts, data theft, manipulation of information, from one group to another, from one state to another. Fake news content thus can play with ideologies to sow mistrust and division within societies, among nations.

#### **4. Fake News Strategy: Bias, Form and Content**

Since fake news is what users are confronted with in their daily endeavour to stay informed and to inform others, we will concentrate on fake news of all the malicious social media uses. On a deeper level, fake news attracts due to the brain structure - *the cognitive biases* - and to the congruent emotional structures. Research has shown that information is decoded based on what one has already experienced or learned, based on his own belief, a process coined as cognitive bias. An example of how cognitive bias works is the one about Frida Sofia, the 12-year-old girl in Mexico. In the context of a very serious earthquake in Mexico, in 2017, there was a reported case of a 12-year-old girl named Frida Sofia who apparently was caught under the rubble of the school she was attending. Mexicans were captivated by watching the news while rescuers were struggling to find the girl even with dog rescue teams. Eventually, the story was reported to have been fake. The question is why the event had been so captivating, why it was so successful and truthful and nobody in the audience questioned the validity while watching. In the end, media stated that Frieda "won the world's hearts, except she didn't exist". The answer lies in the collective desire to help that they all felt and in the knowledge of the seriousness of such cases, gained from previous similar earthquake experiences, some years before. The Mexicans vigilance was taken over by biases. These stories, though fake, were successful and appealing due to the *confirmation bias*. With confirmation bias one's brain looks for information that confirms what one already believes to be true. Also, the *implicit bias* that refers to stereotypes which affect

understanding, actions, decisions, in an unconscious manner is what helps people believe the false information and eliminate what is contrary to their own beliefs. These are what malicious actors rely on, when they build effective fake news, knowing that people's attraction towards the news is driven with famous names and places that are easily accepted by the source, a few real facts from the temporal proximity, while many new contextualized sources provide legitimacy, appealing to eyewitnesses, to psychological and emotional investments. Other examples of successful fake news include the ones inoculated through the COVID-19 pandemic anxiety. For example, most of Russia's generated fake stories at these times were about western countries that apparently are responsible for all the phenomenon, in stories that mention the virus was created as a weapon of mass destruction, as a biological weapon which fell from the sky or which was fabricated in the laboratory, either for the US, for Europe or both, a theme also popular in Chinese disinformation. All the social media posts spread this in more than 1000 shares, according to the EU Observer (2020), a figure that shows the proliferation of the news. The credibility is built on the older cognitive bias with reference to the China-US-Russia relationships and their image in the world political arena. Most people get their information from social media. It has been shown that adults consume 65% of the news from inordinate options- blogs, texts, YouTube, GIFs, emoji, sites, notifications and from the opinions of family members. Throughout these instruments, people consume also propaganda, disinformation, conspiracy theory, click-bait, satire, bias challenge - all manifestations of the same *fake news*, spread within mainstream news to influence others' opinions or to confirm their own opinion through consumed information (Lagarde & Hudgins, 2018).

In an endeavour to raise awareness on how to judge and separate the informative core from the deceptive one, European Association for Viewers Interests, an NGO in Brussels that fights for spreading media literacy across Europe has identified ten forms that fake news may take, irrespective of the demographic features:

- 1) *Propaganda* has been adopted by governments, corporations and NGOs to manage attitudes, values and knowledge. It appeals to emotion and it can be harmful or beneficial. It is a mind game that plays with people's feelings, it juggles with fears and prejudices and is built on individuals' personal relevance.
- 2) *Click-bait* is an attractive sensational title, meant to appeal. It is often misleading, as the content does not always reflect it. However, it brings revenue for advertising or profit from visitors' sites. Click-bait titles often contain famous or familiar names, to attract interest.
- 3) *Sponsored content* is actually advertising made to look like an editorial and is used where there is a potential conflict of interest for the news organizations. The users might not identify the content as being advertorial, if it is not signalled properly and it might thus pass as informative material.
- 4) *Satire* is seen in social or humoristic commentaries which vary a lot in quality. The intended meaning may not be apparent and it can harm those people who misunderstand the message and take it as being real.

- 5) *Errors* are fake news generated by omission: in very well managed news outlets sometimes mistakes are made, and those mistakes can bring prejudice to the brand, prejudices that can be solved through correcting messages published afterwards
- 6) *Bias or partisan information* is ideological and includes interpretations of facts that can pretend to be impartial. It privileges facts that conform to some of the narratives while it ignores others, it is built on emotional and passionate language
- 7) *Conspiracy theory*- tries to explain complex realities as an answer to fear our uncertainties. Proofs that demolish conspiracies are regarded as extra proofs for that very theory, since they reject experts and authorities.
- 8) *Pseudo-science* as a type of fake news based on denials of popular, viral topics like vaccines, miraculous cures, climatic changes, erroneously presented scientific changes with fake or augmented claims, often contradicting the experts.
- 9) *Disinformation* includes a mix of fake, partially fake or factual information, with an intention to inform. Sometimes, authors may not be aware that they used fake information with misleading titles, fake attributions and doctored content
- 10) In a nutshell, fake news is built with a fully changed content with the intention to dis-inform. To spread the content, guerrilla marketing tactics (information presented in a shocking, unusual manner), bots (software agents) or counterfeited comments and brands are used, motivated by the desire for power, politics or both. Regarded from the content point of view, one is faced with *misleading content*, on the one hand, defined as such by the content which is different from what titles or subtitles present, or with *doctored content*, faked graphs, statistics, video content or photos. In the same context, one may have to differentiate the *fake attributions*, authentic images, videos or quotes attributed to wrong events or people or *counterfeit media*, websites or Twitter accounts that impersonate a popular brand or a person, without any connection with reality. Not all of the communication products mentioned above have the same impact or the same motivation power for the targeted user. In the same context, authors who identified all these malicious communication problems have measured the impact and the motivation power. Thus, a strong impact is connected to fake news, to misinformation, conspiracy theory and pseudoscience, medium impact characterizes the partisan products while the click-bait, satire, sponsored content and errors have low impact. Propaganda has been considered to be neutral. On the other hand, looking at their power to motivate, propaganda provides motivation for political power, the click bait has a financial and entertainment motivator, satire is meant to entertain, error – to dis-inform, conspiracy theory is aimed at disinformation while pseudo-science is a driver for power and finance. Disinformation motivates people to build mistrust while fake news as an instrument is used for power and finance.

## 5. Fake News Prevention Strategies

All of the misleading, deceitful content runs based on strategies that, once noticed, can allow individuals to become literate users of informative content, safe from becoming subject to manipulation with any aim.

The malicious content is repetitive and it has an exaggerate dissemination of the same subject, often including real facts to make it difficult to differentiate between truth and lie and it always comes in the form of media products - articles, video, multimedia that very well resemble mainstream media. Sometimes, the materials may be endorsed by testimonials which are fabricated, taken out of context for legitimacy. In terms of message content, they are based on fear or propaganda to fire anxieties already existent or speculate human trend to believe the information that confirms our biases and that stimulates the emotional side.

Beyond mastering the strategies used for the malicious content dissemination inside social media, guardians of truth in the persons of strategic communication specialists have considered a set of questions that provide a better control of the media product prepared for consumption. The questions have been gathered in a place with some control rules, together being an instrument for accurate perception of what reality defines nowadays. Among the most relevant, we mention locating information (which site distributed it), critically evaluating it and communicating it through triangulation.

Beside all of the above, the CRAAP test developed by Lagarde & Hudgins (2019) activates media competences and helps check for fake information. The test comprises a series of questions applied to the piece of news to check if the information in focus is true or meant to dis-inform, not appealing to sense but to sensibility. The authors consider CRAAP to stand for current information, relevant for the reader and for the searched context, the accuracy of data and the goal are elements to be looked at along with the authorship of the material, if the source is credible. In what follows, the adjacent test questions for each of the already mentioned elements are present below:

- 1) *Current* – when was the information published? Are the references new? Is actuality important?
- 2) *Relevant*- is the information connected to the story and tied to the topic? What audience is it written for? Is the proper approach taken?
- 3) *Authorship* – who is the author? Is he qualified? What is the author's status? Is he evaluated? Is the url (link) relevant?
- 4) *Accuracy* - where does the information come from? Are there references? Are there errors? Are there shortened links?
- 5) *Plan* – what is the aim of this piece of information? Is it made for advertising? Is it meant for academic purpose? Is it an opinion? Is it bias?

## 6. Conclusions

Data is now the vehicle people use to influence, since the digital informative environment holds enormous amounts of it and since it is cheap to influence even the most trained individuals by just appealing to social media. For this reason, communication and education build awareness over the malicious use of social media in order to allow individuals to identify and analyse instruments of unconventional warfare and find ways to protect themselves from disinformation. What is more, is that at present facts do not matter anymore, it is perception that has become most important, since it is targeted in manipulating people to gain power and control just by speculating



the faded difference between fact and opinion and between harmful content and what is meant as accurate news, in the abundant information, encountered in social media. In this respect, communicators have already produced rules of conduct and strategies to highlight these differences along with raising awareness over all the tactics used in the process of sowing mistrust and polarization, to divide people and weaken society's confidence in state institutions. Examples of content attacking people's cognition and behaviour in a desire to influence are inordinate, especially in social media, most often with influencers and celebrities that replicate malicious data chunk, lost among the bots' and trolls' contribution. For all these, a thorough awareness of the forms fake news takes and how it replicates as well as what attitude is best in identifying, avoiding and counter-reacting against it are important for all the social media users.

## References

- Bastos, M.T. & Mercea, D.(2019). The Brexit Botnet and User-Generated Hyperpartisan News. *Social Science Computer Review*, 37(1), 38–54, <https://doi.org/10.1177/0894439317734157>.
- Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*, 21(11), <https://doi.org/10.5210/fm.v21i11.7090>
- Cambridge Dictionary, (n.d.) Impersonate. In *Cambridge Dictionary.org*, retrieved June,10, 2020,<https://dictionary.cambridge.org/dictionary/english/impersonate>
- European Association for Viewer's Interest, *Beyond Fake News- 10 Types of Misleading News*, <https://eavi.eu/> (accessed April 11,2020)
- Ferara, E. (2017). Disinformation and social bot operations in the run up to the 2017 presidential election. *Computer Science, First Monday*, 22(8) Cornell Univ.
- Freberg, K. (2019). *Social media for strategic communication*. Sage Publications
- Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide*. New York: New York University Press.
- Jenkins, H. (2007). Transmedia Storytelling. Confessions of an Aca-fan. <http://www.henryjenkins.org>
- Lagarde, J. & Hudgins, D. (2018). *Fact vs Fiction*. Oregon, USA: International Society for Technology in Education.
- Malinen, S. (2015). "Unsociability" as Boundary Regulation on Social Network Sites. Paper presented at the Twenty-Third European Conference on Information Systems (ECIS), Münster, Germany, 26–29 May. [http://aisel.aisnet.org/ecis2015\\_cr/128](http://aisel.aisnet.org/ecis2015_cr/128) (14) (PDF) *Controlling social media flow: Avoiding unwanted publication*. Available from: <https://www.researchgate.net/publication/312065248> [Controlling social media flow](https://www.researchgate.net/publication/312065248) [Avoiding unwanted publication](https://www.researchgate.net/publication/312065248) [accessed Mar 29 2020].
- Robinson, O. (2018), malicious use of social media: case studies from BBC monitoring, <https://www.stratcomcoe.org/malicious-use-social-media-case-studies-bbc-monitoring> retrieved April, 13, 2020
- Rettman, A. (2020, March, 27). Russia's top coronavirus "fake news" stories, *EU Observer*, <https://euobserver.com/coronavirus/147905> , accessed at April 12, 2020.

- Robertson, M. (2010). "A Social approach to security: Using social networks to help detect malicious web content" Thesis. Rochester Institute of Technology. Accessed <https://scholarworks.rit.edu/cgi/viewcontent.cgi?article=1821&context=theses>
- Scolari, C. A. (2009). Transmedia Storytelling: Implicit Consumers, Narrative Worlds, and Branding in Contemporary Media Production. *International Journal of Communication*, 3, 586-606.
- Stenger, T., & Coutant, A. (2010). Vers un management des «112amis» sur les réseaux socio-numériques? Usage et appropriation sur Facebook, Skyrock et MySpace», *XVe colloque de l'Association Information et Management*, La Rochelle, France, May (pp.1-21). <http://hal.archives-ouvertes.fr/hal-00525841/> (14) (PDF) *Controlling social media flow: Avoiding unwanted publication*. <https://www.researchgate.net/publication/312065248> *Controlling social media flow* [Controlling social media flow](https://www.researchgate.net/publication/312065248) [Avoiding unwanted publication](https://www.researchgate.net/publication/312065248) [accessed Mar 29 2020].
- Whitmore, A., Agarwal, A. & Da Xu, L. (2014). The Internet of Things—A survey of topics and trends. *Information Systems Frontiers*, 17(2), pp.261-274.
- Wilson, T. (2004). *The playful audience: from talk show viewers to Internet users*. Cresskill, NJ: Hampton Press.