# Ranked by Truth Metrics: A New Communication Method Approach, on Crowd-Sourced Fact-Checking Platforms for Journalistic and Social Media Content

Evangelos Lampou[1], Nikos Antonopoulos[2]

[1]Department of Digital Media and Communication, Ionian University, Greece

[2]New Media Communication and Usability Lab, Ionian University, Greece

Correspondence: Nikos Antonopoulos, Department of Digital Media and Communication, Ionian University, Greece.

## Abstract

Fake news, misinformation, and non-true stories create a definite threat to the world's public sphere. Fake news contaminates democracy by blurring the sight and the vision, or by altering the beliefs of citizens on simple everyday matters but also on significant matters such as vaccination, politics, social issues, or public health. Lots of efforts have been conducted in order to tackle the phenomenon. Fact-checking platforms consist of a major step in this issue. Certain cases of fact-checking platforms worldwide seem to work properly and fulfill their strategic goals, although functional and other issues might emerge. This study comes to take the fact-checking platform evolution one step beyond by proposing a new communication model for fake news detection and busting. The proposed model's blueprint is based on the Greek "Ellinika Hoaxes" fact-checking platform with some critical reinforcements: More extensive use of crowdsourcing strategies for detecting and busting non-true stories with the aid of AI chatbots in order not only to bust non-true stories but also to rank news outlets, writers, social media personas and journalists for their credibility. This way, serious news outlets, journalists, and media professionals can build their trust and be ranked for the credibility of their services for a more trustful and democratic public sphere.

**Keywords:** crowdsourcing, fake news, fact checking, AI chatbots, media outlets

## 1. Introduction

The term "fake news", refers to false or misleading information that appears as real news. It aims to deceive or mislead people. Fake news comes in many forms, such as clickbait (misleading headlines), disinformation (with malicious intention to mislead the public), misinformation (false information regardless of the motive behind), hoax, parody, satire, rumor, deceptive news and other forms (Raza & Ding, 2022). Fake news can have detrimental effects on individuals and society. First, people may be misled by fake news and accept false beliefs. Second, fake news could change the way people respond to true news. Third, the wide propagation of fake news could break the trustworthiness of the entire news ecosystem (Shu et al., 2017).

Different research approaches indicated that about two-thirds of US, adults experience fake news as a source of confusion and that nearly 25% of them report having shared a fake news story on social media. According to a similar Swedish study of adults aged 18 and older, 4 in 10 reported encountering news on the internet that they believed to be untrue each week. In the same study, 8 in 10 reported believing that false news causes confusion. Others have found that belief in misperceptions is prevalent and plays a significant role in various political and societal issues. (Brandtzaeg et al., 2017). Social media platforms have become where people share the urge to stay connected and express themselves by discussing the news, giving opinions about politics, and watching movies or products, but, unluckily, this online data can be fast and quickly disseminated, and it is not so hard to manipulate with it. Fake news and misleading information have become apparent today, especially in moments of crisis (Jnoub et al., 2021).

Thus, it is important to detect fake news on social media. Fake news is intentionally written to mislead citizens or consumers, which makes it nontrivial to detect simply based on news content. To build an effective and practical fake news detection system, it is natural and necessary to explore auxiliary information from different perspectives (Shu et al., 2019).

Greater public awareness is essential for improving societal resilience against the threat that disinformation poses. The starting point is a better understanding of the sources of disinformation and of the intentions, tools and objectives behind disinformation, but also of our own vulnerability. Increasing skills, knowledge, awareness of the scope of the problem of misinformation, and commitments to fight misinformation are avenues for building resilience (Rodríguez-Pérez & Canel, 2023). A vast and growing literature on the power of the "wisdom of the crowds," or collective intelligence, provides evidence for the efficacy of aggregating non-expert responses in domains ranging from forecasting political and economic events to predicting sporting outcomes and weather trends. Recently, the utility of crowdsourcing has also gained momentum in a number of scientific fields, including medicine, geology, and astronomy (Godel et al., 2021).

One of the most interesting developments is the creation of general-purpose markets for crowdsourcing diverse tasks. For example, in Amazon's Mechanical Turk (MTurk), tasks range from labeling images with keywords to judging the relevance of search results to transcribing podcasts. Such "micro-task" markets typically involve short tasks (ranging from a few seconds to a few minutes) which users self-select and complete for monetary gain (typically from 1-10 cents per task). These markets represent the potential for accomplishing work in a fraction of the time and money required by more traditional methods (Kittur et al., 2011). Co-creation is a subtype of crowdsourcing that refers to a dynamic, interactive process wherein journalists and readers collaborate on a journalistic product in a structured, systematic manner. Co-creation is typically a longer-term process than crowdsourcing, which tends to take place in short-term campaigns. Another subtype is crowdfunding, in which the crowdsourcer asks the crowd for financial contributions. In an open call, the crowd is invited to participate in an online task by submitting information, knowledge, or talent. In crowdsourced journalism, the crowdsourcer asks participants to contribute to an open task online by sharing their knowledge or completing a variety of online tasks. The crowdsourcer can be a journalist, a group of journalists, or a media organization (Aitamurto, 2019). Considering all the above crowdsourcing -taking into account its weaknesses- could be a significant strategy in fighting fake news and misinformation.

## 2. Literature Review

How can misinformation be tackled? On the one hand, some studies suggest that fact-checking shows promising effects depending on the timing and the kinds of labels used. On the other hand, however, fact-checking alone is deemed insufficient to correct misperceptions or, in some cases, is even shown to be counterproductive such that it can reinforce inaccurate beliefs. In a similar vein, media literacy courses and general warnings about the presence of misinformation can also have unintended spillover-effects, making people critical toward not only misinformation but also toward factually accurate information. Furthermore, recent research looking beyond misperceptions as an outcome finds that news media's attention to misinformation decreases trust in science and politics (Hoes et al., 2023).

Cross-validation procedures and initiatives have emerged as the most useful practices to separate fact from fiction. In this direction, platforms like Facebook have taken steps to help citizens access verified information, connecting all them to related procedures and repositories. Since, fact-checkers have played an important role in this battle by critically evaluating claims, thus informing citizens about their integrity and credibility (Katsaounidou et al., 2019).

Broadly speaking, manual fact-checking can be divided into (I) expert-based and (II) crowd-sourced fact-checking. Expert-based fact-checking relies on domain-experts as fact-checkers to verify the given news contents. Expert-based fact-checking is often conducted by a small group of highly credible fact-checkers, is easy to manage, and leads to highly accurate results, but is costly and poorly scales with the increase in the volume of the to-be-checked news contents. Recently, many websites have emerged to allow expert-based fact-checking better serve the public. We list and provide details on the well-known websites in Table 1. Some websites provide further information, for instance, PolitiFact provides "the PolitiFact scorecard", which presents statistics on the authenticity distribution of all the statements related to a specific topic. This information can provide the ground truth on the credibility of a topic, and help identify check-worthy topics that require further scrutiny for verification. Another example is HoaxSlayer, which is different from most fact-checking websites that focus on information authenticity because it further classifies the articles and messages into, e.g., hoaxes, spams, and fake news. Though the website does not provide clear definitions for these categories, its information can be potentially exploited as ground-truth for comparative studies of fake news. In addition to the list provided here, a comprehensive list of fact-checking websites is provided by Reporters Lab at Duke University, where over two hundred fact-checking websites across countries and languages are listed. Generally, these expert-based fact-checking websites can provide ground-truth for the detection of fake news, in particular, under the broad definition. Among these websites PolitiFact and GossipCop have supported the development of fake news datasets that are publicly available (Zhou & Zafarani, 2020).

Table 1. Major Manual fact checking platforms (Zhou & Zafarani, 2020)

| Platform | Topics Covered | Content Analyzed | Assessment Labels |
|---|---|---|---|
| PolitiFact | American politics | Statements | True; Mostly true; Half true; Mostly false; False; Pants on fire |
| The Washington Post Fact Checker | American politics | Statements and claims | One pinocchio; Two pinocchio; Three pinocchio; Four pinocchio; The Geppetto checkmark; An upside-down Pinocchio; Verdict pending |
| FactCheck | American politics | TV ads, debates, speeches, interviews, and news | True; No evidence; False |
| Snopes | Politics and other social and topical issues | News articles and videos | True; Mostly true; Mixture; Mostly false; False; Unproven; Outdated; Miscaptioned; Correct attribution; Misattributed; Scam; Legend |
| TruthOrFiction | Politics, religion, nature, aviation, food, medical, etc. | Email rumors | Truth; Fiction; etc. |
| FullFact | Economy, health, education, crime, immigration, law | Articles | Ambiguity (no clear labels) |
| HoaxSlayer | Ambiguity | Articles and messages | Hoaxes, scams, malware, bogus warning, fake news, misleading, true, humour, spams, etc. |
| GossipCop | Hollywood and celebrities | Articles | 0-10 scale, where 0 indicates completely fake news and 10 indicates completely true news |
| Ellinika Hoaxes | Politics, religion, nature, aviation, food, medical, etc. | TV ads, debates, speeches, interviews, news, etc | Misinformation, Pseudoscience, Scams, Fake News and Hoaxes, Mixing of Facts and Falsifications, Conspiracy Theories, Myths & History |

Some initiatives include Facebook's fact-checking program and Climate Feedback's use of domain experts. However, expert feedback is hard to scale. Other initiatives such as TruthSquad, FactcheckEU, and WikiTribune have pursued a lower-barrier crowdsourced approach, which sometimes run into issues with quality; workarounds include final judgments by experts or delegating primary research to experts and secondary tasks to the crowd. Efforts to automate fact-checking still require human judgment and advances in understanding the crowd labeling of data […]. It is still unclear whether crowdsourcing can be an effective strategy for assessing misinformation at larger scales. Partly this has to do with the limits of crowds on certain topics. It is accepted that collective wisdom can be better than an individual's judgment, including those of individual experts. However, there are situations in which the collective is a lot worse because they do not have enough relevant information, suggesting a baseline expertise in the crowd is necessary (Bhuiyan et al., 2020).

*2.1 Where to Track Fake News?*

Non-true stories, fake news, and misinformation penetrate many different media outlets, even legacy and historic media. Brand qualities such as the trustworthiness of traditional media outlets might be under question. This mainly happens because of the time pressure to editors and journalists, which leads them lots of times not to cross-check their material.

Fake news debases the truth and causes "confusion of fantasy and reality" in the minds of the audience. There are three significant factors that characterize the current brand of fake news. First, the ease by which people can create contents; second, the distribution patterns across new and social media, and third, the political economy of the online domain which

enables and incentivizes the creation of these forms of news. Anyone with internet access can effectively produce and distribute contents of any kind of quality and that "people can avail of the various functionalities of computer software such as Photoshop and create highly believable content (Pate et al., 2019).

Though legacy media really seems not to publish non-true stories intentionally, compared to other outlets, there is no absolute guarantee that they are completely clear from non-true stories. Consequently, the popularity of content providers does not always guarantee the quality or the accuracy of the information given. It has also to be understood that popularity and legacy of the media outlets do not always coincide. Social media provides the ability for fake news to spread fast and have made the situation even more complicated. Fact-checking platforms seem like a fair and promising solution to tackle the fake news phenomenon, proving in action their effectiveness (Lamprou et al., 2021)

Crowd or participatory journalism refers to professional news production with active audience participation and inclusion of citizen generated content in professional journalism. This form of content intends to supplement the professional news production processes and differs from the autonomous content published on social media. Media organizations are embracing participatory journalism to overcome the crisis of traditional journalism falling into a cycle of lower audiences, revenue, content quality, and credibility […].The advantage of crowd journalism is its potential to attract new users, foster transparency and credibility in the news generated, build trust regarding its content, and increase web traffic. Its conditions and requirements include decentralization, collaboration, and (non-proprietary) democratization, attracting more citizens to contribute to news production actively. Crowd journalism enables the publication and dissemination of news that otherwise would not reach the public and potentially impacts all information areas (Lima et al., 2023).

*2.2 Automated Journalism and Artificial Intelligence*

Automated journalism, also known as algorithmic journalism, computational journalism and robot journalism, refers to the application of computer programs— namely, algorithms—to news work, with the aim to organize, interpret and present news pieces from structured data sets. An algorithm is a set of specific instructions programmed to perform a given task and solve a precise problem. Within the vast research field of Data Science, machine learning is the branch of artificial intelligence dedicated to the computational ability of automating data analysis and model building. In simple terms, data scientists develop methodologies and techniques that allow computer programs to learn from data, hence the name, which have the potential to be applied to highly time and labor-intensive tasks for humans. In this precise sense, automated journalism corresponds to the application of algorithmic news judgment and the automation of specific journalistic tasks related to news reporting, writing, curation, data analysis and online social platforms' dissemination (Pérez-Seijo & Vicente, 2022).

Artificial intelligence (AI), defined as "a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation", can be divided into three types: analytical, human- inspired, and humanized. Analytical AI has characteristics consistent with cognitive intelligence only and such a system could learn the campaign platforms of various parties and respond to questions from citizens with respect to the contents. Human-inspired AI has elements of cognitive and emotional intelligence, i.e., understanding human emotions, in addition to cognitive elements, and uses these in its decision making. Such a system could use facial expression to detect when a citizen appears to have problems understanding a party's platform, and can providing him or her with more information. Humanized AI exhibits characteristics of all types of competencies (i.e., cognitive, emotional, and social intelligence), and is able to be self-conscious and self-aware in interactions with others (Kaplan, 2020). Chatbots or conversational bots can be defined as software programs (or even, in some cases, hardware) that are able to respond to text or even voice messages. Those programs attempt to simulate responses of a human being, thus act as natural language user interfaces for data and service providers. They can be programmed to respond differently depending on the existence of certain keyword or they can be more sophisticated and employ machine learning techniques to adapt their responses based on the included words in the queries (Veglis & Maniou, 2019).

Technology plays a massive role in the industry and daily chores. It serves a variety of purposes and is applied in a different way in different parts of the world. Recently, the public has been fantasized by Artificial Intelligence. Artificial Intelligence simulates the cognitive abilities of a human. To be more precise and closely related to humans, the AI Chatbots are now replacing human responses with this software. A Chatbot is a computerized program that acts like a colloquist between the human and the bot, a virtual assistant that has become exceptionally popular in recent years mainly due to dramatic improvements in the areas like artificial intelligence, machine learning and other underlying technologies such as neural networks and natural language processing. These chatbots effectively communicate with any human being using interactive queries. Recently, there's been a massive increase in many cloud-based chatting bot services which have been made available for the development and improvement of the chatbot sector (Gupta & Hathwar, 2020).

AI chatbot technologies have been accused for being a potential hazard for the public sphere and mankind in general. Tech companies are moving too fast in rolling out powerful artificial intelligence technology that could one day outsmart

humans. That's the conclusion of a group of prominent computer scientists and other tech industry notables such as Elon Musk and Apple co-founder Steve Wozniak who called for a 6-month pause to consider the risks. Their petition published on 20/03/23 is a response to San Francisco startup OpenAI's recent release of GPT-4, a more advanced successor to its widely-used AI chatbot ChatGPT that helped spark a race among tech giants Microsoft and Google to unveil similar applications. The letter warns that AI systems with "human-competitive intelligence can pose profound risks to society and humanity" — from flooding the internet with disinformation and automating away jobs to more catastrophic future risks out of the realms of science fiction. It says "recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control." "We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4," the letter says. "This pause should be public and verifiable, and include all key actors. If such a pause cannot be enacted quickly, governments should step in and institute a moratorium (O'brien, 2023).

### 2.3 Automated Fake News Detector Initiatives

The "Fake News Detection Task," which assesses the truthfulness of a certain piece of news from news content or social media posts, has been performed by many researchers to save working hours and make the process automatic. In recent years, with the development of deep learning models, many models have been proposed to achieve high detection performance. A classical fake news detection model learns the textual style of fake news and then classifies them as fake on the basis of input data, such as information of the news content or social media posts, including the news […] It is difficult to detect fake news generated in the last few years using only textual information because the representation of fake news has become more diverse and complex. Recent fake news detection models leverage textual information as well as rich contexts such as news publishers, user information, temporal information, and network information. Additionally, international competitions such as CheckThat! and Constraint intensified the activities in proposing a novel fake news detection model. Dataset construction, which provides resources for such research, is also an essential step in detecting fake news. This section introduces the fake news detection dataset from the following two perspectives;

• News articles: Datasets that are utilized to detect fake news mainly from the body of the news article. The style of each news article is an important feature for detection.

• Social Media Posts: Datasets that are utilized to detect fake news mainly from social media posts related to each news. User information and network information in social media, in addition to text in social media posts, are important features (Murayama, 2021).

Another notable fake news detector endeavor is Check-It. According to its creators, if current fake news detection and deployment approaches are analyzed, two issues are raised: first, the need to consider and combine more signals in the identification process, in order to boost its overall effectiveness. Notably, prior approaches utilize a single input signal (flag lists, fact-checks, article content, or social networks). To this end, however, it seems appropriate to come up with more effective signals and approaches for combining them efficiently. The second issue is related to the preservation of end-user privacy when assessing visited pages, by not revealing the user's identity and browsing history to any third-party services, in compliance with the EU's GDPR policy. To address these issues, Check-It was designed and implemented, and a fake news identification system was developed as a browser plugin. Check-It bundles together a series of diverse signals, including flag lists, similarity matching, and Artificial Intelligence (AI) techniques, making it able to calculate the credibility of a piece of news and successfully warn the reader, whilst securing his/her privacy (GDPR compliant) by working locally on the browser without the need for external communication (Paschalides et al., 2021).

### 2.4 Public Sphere and Fact-Checking in Greece

According to the latest Reuters Institute for the Study of journalism Report (2021), the media market in Greece is characterized by digital fragmentation, lack of trust in news, a politically polarized press, and one of the highest uses of social media for news. The majority of Greece's audience (89%) gets the news through online media (including social media) and only 22% replied getting the news by print outlets. 37% of the respondents said that they share news via social, messaging or email. As in most other countries trust in news increased, by four percentage points in Greece (32%). Despite the increase, Greece is still well behind most other countries in trust (38th place out of 42 countries). According to the Standard Eurobarometer 94 (Winter 2020 - 2021), the proportion of respondents reporting they watch TV on a TV set every day or almost every day is estimated at 73%, the proportion of people listening to the radio every day or almost every day has risen particularly and is estimated at 42%. Levels of trust in television were lowest in Greece (25%), whereas trust in the Internet has grown notably (54%, +12 percentage points). 60% of the respondents consider that the national media does not provide trustworthy information. Finally, regarding public media, 52% of the respondents believe that the statement that public service media are free from political pressure does "not at all" reflect the situation in their country (Papadopoulou, 2022).

"Ellinika Hoaxes" fact-checking platform commenced in Greece as a blog as an idea of Thodoris Daniilidis, investigating

the news circulating on the Greek internet and highlighting those that are not true. Generally speaking, Laura Bononcini, Director of Public Policy of Facebook in Southeast Europe, officially states on May 2, 2019, the beginning of a partnership with "Ellinika Hoaxes". The platform does not produce original news stories and it is not a journalistic media outlet. It is merely a fact-checking website funded by Facebook, combining fact-checkers employment and crowdsourcing strategies in order to bust fake news that emerges in the Greek public sphere. "Ellinika Hoaxes" uses a team of professional fact-checkers who run the fact-checking procedure based on information that comes from the crowd through an open call: "Ellinika Hoaxes encourages readers to participate in the fight against fake news. For this reason, we are always open to suggestions, remarks, corrections, submission of topics for research, etc. Your participation through the submission of proposals is one of the basic rules for choosing our topics" (Ellinika Hoaxes, 2023). According to Lamprou & Antonopoulos (2020), all major crowdsourcing strategies of Howe's are used (Crowd creation, Crowdvoting, Crowdwisdom, except Crowdfunding. The crowd proposes potential non-true stories and fake news cases through the "Ellinika Hoaxes" website or official Facebook page. Then, professional fact-checkers of the platform take action.

## 3. Findings: A New Fact-Checking Communication Model Blueprint

Considering the advantages and drawbacks of "Ellinika Hoaxes" and deep learning and AI potentials, this paper aims to propose an enhanced fact-checking communication model based on fundamental crowdsourcing principals aided by Artificial Intelligence chatbots. In other words, monitoring the public sphere for the detection and busting of fake news, misinformation and disinformation, through specialized fact checking techniques seems to work for the established Greek platform "Ellinika Hoaxes. Though, criticism has been released against the platform by some Greek news media websites suggesting that there are biased fact-checking cases and not always accurate in their judgments.

Taking into account the literature review and the major fact-checking platforms functions worldwide and in Greece specifically, researchers aim to present a composite new crowdsourcing-based fact-checking model. The methodology applied was to make the assumption, that we can use established and proven parts from tested and successful platforms referred to in the literature review combined with new innovating elements such as artificial intelligence chatbots and different crowdsourcing strategies in order to compose a new communication model blueprint. The model is presented as a blueprint and explained thoroughly in three stages and is presented in 3 different schemes. Each stage is described separately in detail. Though, as it is referred to in the limitations section, the model blueprint has to be tested and evaluated in action in order for the researchers to evaluate its strengths and weaknesses and make potential improvements in order for the new platform to function properly.

The proposed platform model blueprint presents an enhanced version of a fact-checking platform, divided in 3 fundamental sections:

### 3.1 Stage 1: Incoming Procedure

The first and most important section of the proposed platform's blueprint is the incoming phase. Here it is to be decided which cases of non-true will be accessed. "Ellinika Hoaxes" fact-checking platform use crowdsourcing to a significant extend in order to decide which fake news to scan. Usually many of the cases are proposed from the crowd (citizens) but also the platforms Internal fact checking teams takes into consideration what is more widespread in the social media outlets. The new platform's model suggests even greater use of crowdsourcing with the users' engagement to be rewarded to a one to 5 stars climax. The users whose proposals are finally selected to the next phase more often will gain ethical reward and acknowledgement (zero to five stars). This acknowledgement may give to some of them reputation end might help them to enter to the professional world of media and journalism.
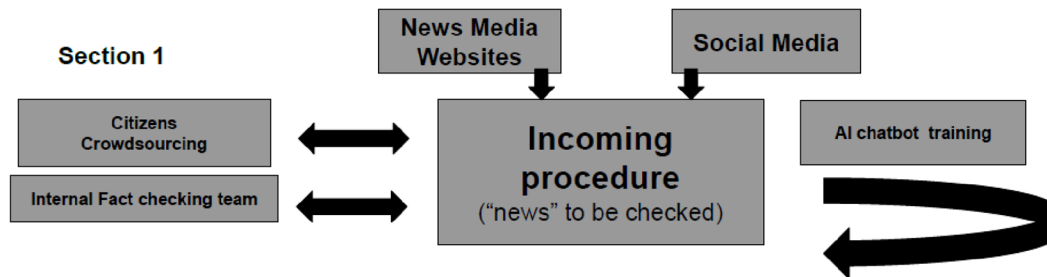


Figure 1. Section 1 Incoming procedure

The platform's internal fact checking team of experts is mainly assigned to run the fact-checking procedure. These are professional fact-checkers who are assigned to decide which of the examined cases are true of fake and to what extent. Though, their judgment on what is finally worth to be examined in collaboration with the auditing team (see next phase) is a crucial supplement to the crowd's contribution. Thus, the incoming material from either social media or news outlets

should be proposed from citizens and the internal fact checking team in combination with the second having the final word on what is to be accessed.

### 3.1.1 Artificial Intelligence Chatbots in the Battle Against Misinformation

AI chatbots might be a threat for public sphere and lead to a huge up rise of non-true stories production and circulation. Recent advancements in natural language generation (NLG) technology have significantly improved the diversity, control, and quality of LLM (Large Language Models)-generated texts. A notable example is OpenAI's ChatGPT, which demonstrates exceptional performance in tasks such as answering questions, composing emails, essays, and codes. However, this newfound capability to produce human-like texts at high efficiency also raises concerns about detecting and preventing misuse of LLMs in tasks such as phishing, disinformation, and academic dishonesty. For instance, many schools banned ChatGPT due to concerns over cheating in assignments, and media outlets have raised the alarm over fake news generated by LLMs. These concerns about the misuse of LLMs have hindered the NLG application in important domains such as media and education (Tang et al., 2023).

But this is not the case here. AI chatbots such as ChatGPT could be used as valuable tools for busting non-true stories getting trained by the platform procedures, professional journalists, and fact checkers. In addition to crowdsourcing strategies and the proposals of the internal fact-checking team, the aid of an AI chatbot to the procedure could be important. The AI chatbot should not be used in order to replace the previously mentioned procedures, but to help tackling the huge volume of cases and incidents to be examined. In all cases AI chatbots should be used under surveillance of the human factor of the platform and only proposing cases to be examined. The final word should fall on the human decision. The AI chatbot can be set and trained by the results and choices made by the platform overall. According to Roose (2023) the procedure is quite simple:

Step 1: Goal setting

A goal is needed for every AI system. It is a procedure which is named an objective function. Most large language models have the same basic objective function: Given a sequence of text, guess what comes next.

Step 2: Data collection

Next, the training data need to be assembled that will teach the AI chatbot how to perform. Ideally, we'll put together a colossally large repository of text, which usually means billions of pages scraped from the internet — like blog posts, tweets, Wikipedia articles and news stories. Before we can feed the data into the model, we need to break it down into units called tokens, which can be words, phrases or even individual characters. Transforming text into bite-size chunks helps a model analyze it more easily.

Step 3: Building the neural network

Once the data is tokenized, we need to assemble the A.I.'s "brain" — a type of system known as a neural network. This is a complex web of interconnected nodes (or "neurons") that process and store information. For our own AI chatbot, we're going to want to use a relatively new type of neural network known as a transformer model. They can analyze multiple pieces of text at the same time, making them faster and more efficient.

Step 4: Training of the neural network

Next, the model will analyze the data, token by token, identifying patterns and relationships. As it learns these patterns, the transformer model sketches a map: an enormously complex mathematical representation of human language. It keeps track of these relationships using numerical values known as parameters. Many of today's best L.L.M.s may have hundreds of billions of parameters or more.

Step 5: Fine-tuning the model

Once a large language model is trained, it needs to be calibrated for a specific job. An AI chatbot used by a hospital might need to understand medical terms, for example. An AI chatbot used for detecting possible non-true stories has to understand all these elements that help it detect them. A lot of these elements can be taught to the AI chatbot by the crowdsourcing procedure.

Step 6: Launch, under supervision

Once the AI chatbot has been trained and fine-tuned, it's ready to use. Although it might seem quite impressive and accurate the AI chatbot has always to be under the supervision of human decisions as companies like Microsoft and Meta have learned the hard way, A.I. systems can be erratic and unpredictable, or even turn creepy and dangerous. This is one of the reasons why in the proposed communication model, human factor retains the core role. According to the findings of Kotonya & Toni, (2020), the emerging work on explainable machine learning in the fact-checking domain shows a great deal of promise despite the particularly challenging nature of the problem. However, there are some limitations of

current methods and task formulations which are highlight in their survey: all existing methods only look to explain one component of the fact-checking pipeline (relation prediction or entailment) and the systems only explain the predictions. In order to garner the most relevant, useful and insightful fact-checking explanations, a holistic and journalistic-informed approach should be taken.

*3.2 Stage 2: Assessment Procedure*

The second section of the platform's blueprint is equally important. Here the final assessment takes place in a combined procedure. The incoming material that has been selected from the crowdsourcing procedure, the internal fact-checking team and the AI chatbot has to be accessed and flagged if necessary. This procedure for the proposed platform blueprint comes through three different processes.

3.2.1 Process 1

Citizens can participate through a call for action gaining ethical acknowledgment for their services and exceptional ranking as users who contribute (zero to five stars). Though, they will have to follow certain procedures in order to flag a story as fake. This reinforces their reputation and fame and might give them the opportunity to enter the world of media and journalism professionals. Citizens should determine if the story or allegation is true or false and they should rate the result from zero to five stars (0= Fake and 5= absolutely true). Also, this will give an opportunity to citizens who know exactly that a story is non-true and they were looking always for ways to communicate with the media in order to make clear with proof that a story is non-true or misleading. However, for a non-true story to be busted, evidence should be presented and uploaded to the platform. Furthermore, crowd wisdom a fundamental element of crowdsourcing according to Howe (2008) is an element that reinforces democracy and gives citizens the opportunity to express themselves. It is estimated that approximately 25% of all cases shall be assessed by citizens. If this goal is not always possible then the remaining cases will be divided equally to journalists' peer review, and to the internal fact-checking team of the proposed platform.
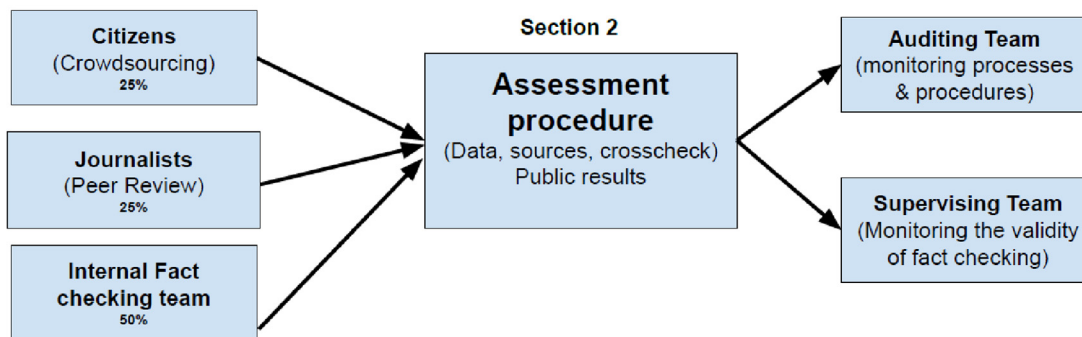


Figure 2. Section 2, Assessment procedure

3.2.2 Process 2

In the proposed platform the second main process of Section 2 is the assessment procedure by journalists as is described in Figure 2: Professional Journalists through an open call shall engage in reviewing the proposed cases of possible non-true stories using their expertise, experience, and their access to journalistic sources and information. However, this procedure has to be carried out by various different journalists and cross-checked in order for the platform to avoid bias. Journalists should determine if the story or allegation is true or false and they should rate the result from zero to five stars (0= Fake and 5= absolutely true). Journalists could assess all cases deriving from social media and blogs but also from legacy media outlets. However, for a non-true story to be busted, evidence should be presented and uploaded to the platform. But why should a news professional or journalist engage in such a procedure? Participating journalists should receive ethical payment for their services and a high score rating for contributing to the platform. Every successful contribution shall increase every participating professional's rating in order for him to gain more fame and trustworthiness in the public sphere. Every journalist and media outlet should receive 2 percent of a star (climax zero to five) for every successful evidence-based fact-checking contribution which will be also crosschecked by the supervising team which will just monitor the validity of fact-checking (figure 2). It is estimated that approximately 25% of all cases shall be assessed by journalist's peer reviewers.

3.2.3 Process 3

The team of professional fact-checkers shall be the main drive train of the assessment procedure. This is a model that works quite efficiently in other major fact-checking platforms such as "Ellinika Hoaxes" (Ellinika Hoaxes, 2023). Existing fact-checking sites and experts are engaged in judging the truthfulness of each piece of news. This is often time-consuming and demands large amounts of manual labor. Nevertheless, it is effective, and human logic and strict methodology seem

to be the best weapon against fake news, giving most of the time satisfactory results. Due to its importance, the internal fact-checking team has to be reinforced with excellent and well-payed professionals whose job is to examine whether a case is fake or not and to rate a scale of truth (0 to 5 stars). Furthermore, as referred to previously the internal fact-checking team shall have the authority to select cases that seem likely to be fake news using their experience, training, and abilities and to put them autonomously in the assessment procedure. When the verdict shall be complete, then after the supervising team's permission the busted non-true story shall be published on the platform with all the evidence in public form, according to the "Ellinika Hoaxes example".

3.2.4 Auditing and Supervising Teams

Of critical importance in this section are the auditing and the supervising teams. The auditing team shall not interfere with the assessment itself and the classification of the potential fake news but shall only monitor the correct implementation of processes and procedures. This team shall be manned with quality assurance managers and experts. On the other hand, the supervising team shall be responsible for the final outcome of the assessment stage, monitoring the validity and accuracy of fact-checking. This team shall be manned with elite fact-checkers, scientists, or renowned journalists and reporters. Both teams are crucial elements for the platform's proper function.

*3.3 Stage 3: Analysis and Ranking*

The third and final stage is data analysis and ranking. All busted cases deriving either from news media websites or from social media and blogs are categorized and stored. News media websites are also ranked on a scale of 1-5 stars with the use of the color red turning to green with yellow intermediate according to true or false. As more cases of fake news a news media website has on the platform, the lower the "truth ranking" it achieves. In order for the platform to display a more analogical and fair truth ranking non-true stories should be divided into the general number of articles published by each website. Furthermore, the non-true story ranking of each individual non-true story, as mentioned above, shall be decisive to the impact it has on the website's overall ranking score. The most crucial and essential element of this process is that it has to be public. All cases of busted non-true stories with the appropriate evidence (as the platform "Ellinika Hoaxes" practices) shall be publicly available even if the news media website has taken it down. The general strategy is for the news media websites primarily and for social media pages and accounts to gain or lose reputation for the truth metrics of their articles and posts. It is a kind of ranking similar to applications such as booking and trip advisor but for news media websites, blogs, and social media pages and accounts based on the concept of truth.
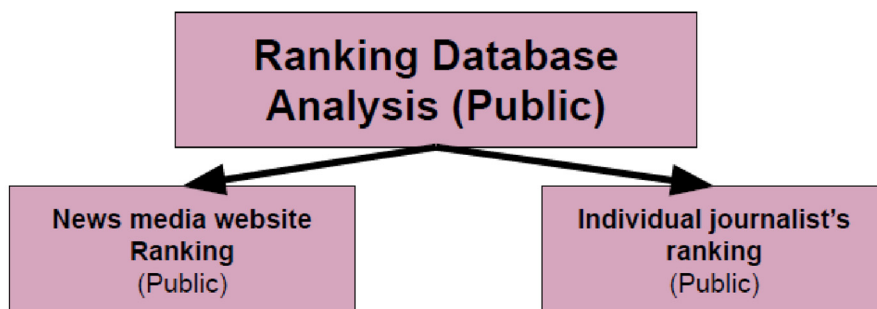


Figure 3. Stage 3, Ranking database analysis

News media websites, blogs and social media accounts shall not be the only ratings output of stage 3. The platform shall give also journalists, reporters, commentators or other media personalities a "truth ranking". This is a way and opportunity for media professionals to build their reputation and stand out in the mediascape according to real metrics and objective goals. The journalist's or writer's "truth ranking" shall be an indication of his or her trustworthiness and all data available for his/her ranking shall be public. Here therefore, comes a question concerning personal data. Journalists, writers and people of the media community should be encouraged to willingly create accounts to the platform in order to gain visibility, fame and to promote their work just as academics do. In any case their work is public and most of the times accessible to everyone, and journalists and media professionals are most of the times public figures. Furthermore, as mentioned before professional journalists can participate in a peer to peer procedure of fact-checking in order to help the platform with their knowledge, experience, and access to journalistic sources.
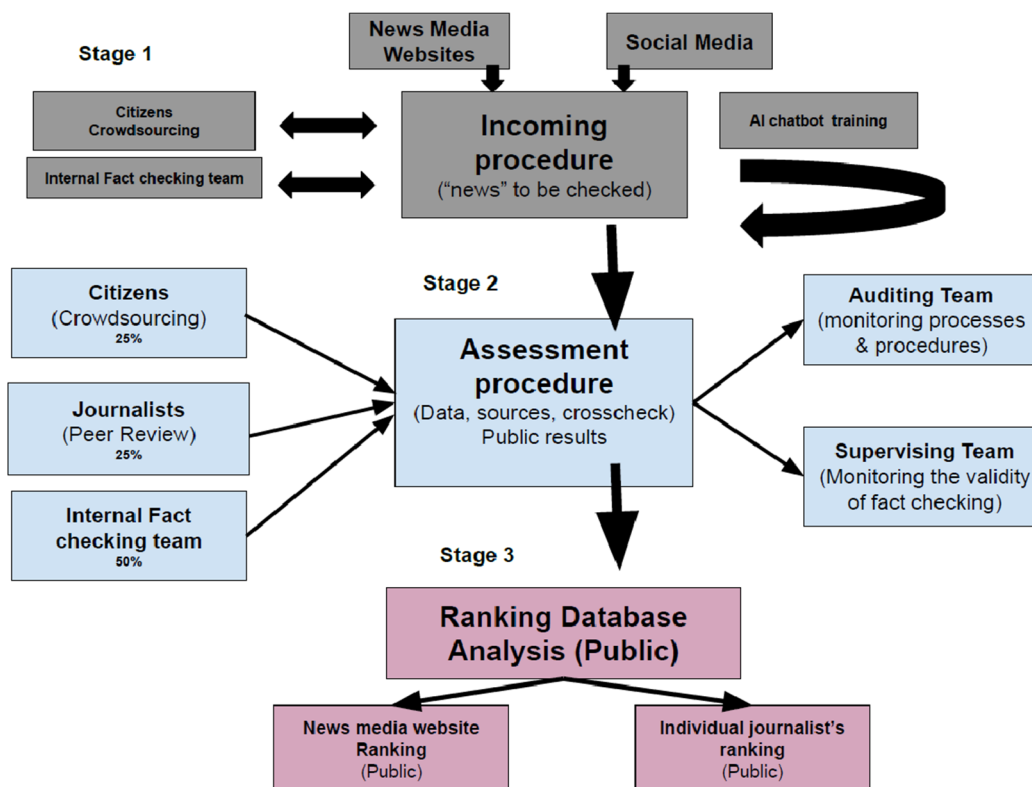
Figure 4. The platform as a whole

## 4. Discussion

There are many potential usages for the proposed communication model and how it could be proven useful both in the public and private sectors. The platform could be used for creating new metrics for journalists, news media websites, and organizations. News media websites, media companies, social media personas and journalists should be acknowledged not only for popularity, circulation, interesting content, and traffic figures but also for their commitment to truth and facts. The proposed platform can provide solid and proven evidence for all participating news media websites and journalists on their commitment to truth. This could make advertisers more skeptical of investing only in popular and high-circulation media outlets, and turn truth commitment into a major factor for high advertising revenue investments. In other words, the proposed platform could change the strategy both of news media websites and major advertisers and create the need for transparency and reliability as a prerequisite. Furthermore, true investigative journalism could be revived. Media organizations shall need quality content and hire high-quality journalists in order to achieve high truth rankings. On the other hand, news media websites and journalists whose ranking is low might have less impact on the public sphere and consequently produce less revenue with their content. The need for higher truth rankings might alter the business model of advertising where the popularity and traffic of free media content lure advertisers. News media websites might also try to avoid easy and not cross-checked content in order to comply with the need for higher truth rankings.

### 4.1 Potential Business Models

The platform could be used as an independent business endeavor as a small or medium-range business or either to begin as a start-up business with a high potential for growth. Revenue for the platform can be derived from sources like recruitment ads and other advertising, both of which are being directed at exactly the audience that those advertisers want to reach. Advertisers might want to positively associate the brand equity of their products and services with a company/organization dedicated to truth. Moreover, it could be specialized to certain aspects of the public sphere or certain categories of news or media such as sports news, political news, entertainment, or social policy news, etc. The ranking of media websites, social media pages, personas, and journalists could lead to universally accepted metrics for truth and transparency all over the world. International or public organizations could also use such a platform in order to possess transparent and broadly accepted evidence for regulating authorities and battling fake news without the state's direct engagement. In any case, the platform proposes a new communication model where the crowd and artificial intelligence chatbots propose, and then, specialized fact-checkers, in combination with citizens and the journalistic community determine whether cases highlighted as suspicious are fake news (partly or completely) or not. Experts should audit and monitor the procedures and processes. When the results are gathered and analyzed, truth rankings for journalists,

news media websites, social media accounts, and personas should be presented publicly. Without a doubt, this platform needs large numbers of specialized personnel in order to function properly. Artificial Intelligence chatbots such as ChatGPT, Bard, and others can be used during the initial procedure of selecting possible cases of fake news or misinformation as described above in the platform's blueprint. Though the chatbots might select cases of possible fake news articles and social media posts, the final decision is human. Artificial intelligence can help in order the model to cope with the vast volume of possible fake news cases. Nevertheless, the process of fact-checking shall be carried out by the human factor.

The implementation of this communication model guarantees that although chatbots and Artificial Intelligence might take over a significant proportion of human jobs, the main core of the decision-making remains human in such a sensitive sector as news credibility. It is essential for democracy that the public sphere remains free from fake news and misinformation, and that citizens shall have objective information on the credibility of news media outlets, journalists, and public figures. In a world economy where consumers use crowdsourcing on various platforms to rate hotels, restaurants, products, and services and take advantage of such valuable information, it seems to be a necessity for citizens to have access to the truth rankings of news media websites, journalists, and of public figures.

*4.2 Regulating Malfunctioning Media Markets*

The value of this new communication model can be of critical importance, especially for countries such as Greece, Portugal, Serbia, Turkey, and similar where the media and news industry might be malfunctioning. Fake news and misinformation/disinformation can be a serious threat to the economy, public interest, and democracy overall. In Greece, the situation of the Journalistic profession, its standards, and protection presents a worrying high risk (74%). Despite the long existence of journalistic unions, as well as their frequent support for journalists whose work rights are being violated, it is quite common sense that they are not effective in guaranteeing editorial independence and/or respect for professional standards. In 2021, the working conditions of journalists in Greece remained generally unsatisfactory. Various media organizations resorted to arbitrary dismissals of journalists or pay cuts (Papadopoulou, 2022). Thus, beyond battling fake news and misinformation it may help the media market regulation in countries where media are mostly state-funded and supported through public advertisement or depend on major help from state-based enterprises or advertisement deriving from banks as it happens in Greece. With the new proposed communication model, the truth ranking of news outlets, social media accounts, journalists, and media people shall make citizens aware to understand the true value of media outlets, social media personas, and accounts, journalists, and news overall. This means greater independence and confidence for the outlets, media people, journalists, and social media personas who serve the truth. The higher truth rankings shall attract more advertising revenue and more support for high-value content and serious outlets from readers and citizens who wish to get in touch with real news and information providing them with financial support and independence.

In addition, the proposed communication model shall help for the logical introduction of artificial intelligence chatbots as a tool against misinformation and not as one who might potentially create fake news or put an end to an industry. This is essential not only for countries that have a malfunctioning media system but for every democratic society that wants to prevent the misusage of AI chatbots as tools of misinformation, fake news, and propaganda. The proposed platform uses artificial intelligence as an ally and not as a foe and this is a crucial element given that many major countries, scholars and think tanks foresee a dark future for journalism and media in general with the introduction of AI chatbots and applications causing even temporary bans for its use (Goujard, 2023). On the contrary, the proposed communication model has a totally positive stance against AI as it considers it a significant element for battling fake news, propaganda, and misinformation while empowering the human factor which remains the king and the decision maker.

## 5. Conclusions

The proposed communication model of this study uses artificial intelligence advantages in favor of fact-checking teams and professional communities. The AI technology is combined with crowdsourcing strategies in order for the platform to track and terminate fake news cases circulating in both media outlets and social media. Despite the rumors that artificial intelligence applications might end lots of professions including journalists and content creators, the proposed model needs high-skilled and experienced personnel in order to function. That means that new business endeavors or even public independent organizations could be triggered aiming to battle fake news and to create a more transparent, clear, and democratic public sphere, by creating new jobs for high-skilled fact-checkers, journalists, quality assurance specialists, programmers, and media people. Moreover, the need for a higher truth ranking, cross-checking, and quality journalism might lead media companies to re-invest in true journalism and investigation and ultimately reshape the industry.

## 6. Limitations – Future Research

The new communication model described in this paper is just a blueprint based on findings of previous research (Lamprou et al. 2021) and on the positive evaluation of the potential of fact-checking platforms in Greece. Nevertheless, the blueprint

has to be experimentally tested and proven on real circumstances on all three stages including incoming procedure, assessment procedure and ranking. The optimum scenario for this includes collaborations with press and journalistic unions, researchers and universities, media companies and the funding of a business plan based on the blueprint of the proposed innovating communication model. Authors plan in the future to proceed in the implementation of the model evaluating all three stages and make improvements and changes if necessary.

**Disclosure Statement**

**Funding**

**References**

Aitamurto, T. (2019). Crowdsourcing in journalism. *Oxford Research Encyclopedia of Communication.* https://doi.org/10.1093/acrefore/9780190228613.013.795

Bhuiyan, M., Zhang, A., Sehat, C., & Mitra, T. (2020) Investigating "who" in the crowdsourcing of news credibility. In: *Computational Journalism Symposium.*

Brandtzaeg, P. B., Følstad, A., & Chaparro Domínguez, M. Á. (2017). How journalists and social media users perceive online fact-checking and Verification Services. *Journalism Practice, 12*(9), 1109-1129. https://doi.org/10.1080/17512786.2017.1363657

Ellinika, H. (2023). Available online: https://www.ellinikahoaxes.gr/ (accessed on 24 March 2023)

Godel, W., Sanderson, Z., Aslett, K., Nagler, J., Bonneau, R., Persily, N., & Tucker, J. (2021). Moderating with the mob: Evaluating the efficacy of real-time crowdsourced fact-checking. *Journal of Online Trust and Safety, 1*(1). https://doi.org/10.54501/jots.v1i1.15

Goujard, C. (2023, March 31). Italian privacy regulator bans Chatgpt. POLITICO. Retrieved April 1, 2023, from https://www.politico.eu/article/italian-privacy-regulator-bans-chatgpt/

Gupta, A., & Hathwar, D. (2020). Introduction to Ai Chatbots. *International Journal of Engineering Research And, V*9(07). https://doi.org/10.17577/ijertv9is070143

Hoes, E., Aitken, B., Zhang, J., Gackowski, T., & Wojcieszak, M. (2023). Prominent Misinformation Interventions Reduce Misperceptions but Increase Skepticism. https://doi.org/10.31234/osf.io/zmpdu

Howe, J. (2008). *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business.* New York: Random House.

Jnoub, N., Brankovic, A., & Klas, W. (2021). Fact-checking reasoning system for fake review detection using answer set programming. *Algorithms, 14*(7), 190. https://doi.org/10.3390/a14070190

Kaplan, A. (2020). *Artificial Intelligence, Social Media, and Fake News: Is This the End of Democracy?* Istanbul University Press: Istanbul, Turkey, p. 149. doi: 10.26650/B/SS07.2020.013.09

Katsaounidou, A., Vrysis, L., Kotsakis, R., Dimoulas, C., & Veglis, A. (2019). Mathe the game: A serious game for education and training in news verification. *Education Sciences, 9*(2), 155. https://doi.org/10.3390/educsci9020155

Kittur, A., Smus, B., Khamkar, S., & Kraut, R. E. (2011). CrowdForge. Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. https://doi.org/10.1145/2047196.2047202

Kotonya, N., & Toni, F. (2020). Explainable automated fact-checking: A survey. *Proceedings of the 28th International Conference on Computational Linguistics.* https://doi.org/10.18653/v1/2020.coling-main.474

Lamprou, E., & Antonopoulos, N. (2020). Fake News, Crowdsourcing and Media Outlets in Greece: Is News Credibility a Matter of Professionalism? Paper presented at 11th Asian Conference on Media, Communication & Film, The Kyoto Conference on Arts, Media &Culture, Kyoto, Japan, November 12–14; pp. 129–42.

Lamprou, E., Antonopoulos, N., Anomeritou, I., & Apostolou, C. (2021). Characteristics of fake news and misinformation in Greece: The rise of new crowdsourcing-based journalistic fact-checking models. *Journalism and Media, 2*(3), 417-439. https://doi.org/10.3390/journalmedia2030025

Lima, I. R., Filipe, V., Marinho, C., Ulisses, A., Chakravorty, A., Hristov, A., … & Prodan, R. (2023). ARTICONF Decentralized Social Media Platform for Democratic Crowd Journalism. https://doi.org/10.21203/rs.3.rs-2695964/v1

Murayama, T. (2021). Dataset of fake news detection and fact verification: A survey. *ACM Comput. Surv., 1*, 1-33.

O'brien, M. (2023, March 29). Musk, scientists call for halt to ai race sparked by Chatgpt. AP NEWS. Retrieved April 1, 2023, from https://apnews.com/article/artificial-intelligence-chatgpt-risks-petition-elon-musk-steve-wozniak-534f0298d6304687ed080a5119a69962

Papadopoulou, L. (2022). Monitoring media pluralism in the digital era – Application of the Media Pluralism Monitor in the European Union, Albania, Montenegro, the Republic of North Macedonia, Serbia & Turkey in the year 2021: country report: Greece, European University Institute. https://data.europa.eu/doi/10.2870/805570

Paschalides, D., Christodoulou, C., Orphanou, K., Andreou, R., Kornilakis, A., Pallis, G., … & Markatos, E. (2021). Check-it: A plugin for detecting fake news on the web. *Online Social Networks and Media, 25*, 100156. https://doi.org/10.1016/j.osnem.2021.100156

Pate, U. A., Gambo, D., & Ibrahim, A. M. (2019). The impact of fake news and the emerging post-truth political era on Nigerian polity: A Review of Literature. *Studies in Media and Communication, 7*(1), 21. https://doi.org/10.11114/smc.v7i1.4238

Pérez-Seijo, S., & Vicente, P. N. (2022). After the hype: How hi-tech is reshaping journalism. *Studies in Big Data,* 41-52. https://doi.org/10.1007/978-3-030-88028-6_4

Raza, S., & Ding, C. (2022). Fake news detection based on news content and social contexts: A Transformer-based approach. *International Journal of Data Science and Analytics, 13*(4), 335-362. https://doi.org/10.1007/s41060-021-00302-z

Rodríguez-Pérez, C., & Canel, M. J. (2023). Exploring European citizens' resilience to misinformation: Media Legitimacy and Media Trust as predictive variables. *Media and Communication, 11*(2), 30-41. https://doi.org/10.17645/mac.v11i2.6317

Roose, K. (2023, March 28). How does chatgpt really work? The New York Times. Retrieved March 30, 2023, from https://www.nytimes.com/2023/03/28/technology/ai-chatbots-chatgpt-bing-bard-llm.html

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media. *ACM SIGKDD Explorations Newsletter, 19*(1), 22-36. https://doi.org/10.1145/3137597.3137600

Shu, K., Wang, S., & Liu, H. (2019). Beyond News Contents. *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*. https://doi.org/10.1145/3289600.3290994

Tang, R., Chuang, Y. N., & Hu, X. (2023). The Science of Detecting LLM-Generated Texts. *arXiv Preprint.* arXiv:2303.07205

Veglis, A., & Maniou, T. A. (2019). Chatbots on the rise: A new narrative in journalism. *Studies in Media and Communication, 7*(1), 1. https://doi.org/10.11114/smc.v7i1.3986

Zhou, X., & Zafarani, R. (2020). A survey of fake news. *ACM Computing Surveys, 53*(5), 1-40. https://doi.org/10.1145/3395046