journalism and media

Article

Disinformation Detection: Developing a Categorical Framework Through Thematic Analysis

Antea Boko

Media and Communication, University North, 48 000 Koprivnica, Croatia; anboko@unin.hr

Abstract: In recent years, disinformation has become a significant problem in the media environment. The topic is therefore increasingly relevant in recent research, and authors approach it in different ways. This research aims to provide an answer to the need for a deeper understanding of how to detect and combat disinformation. The primary purpose of this research is to identify and systematize key categories that enable the detection of disinformation, providing a solid framework for combating this ubiquitous challenge. The qualitative method of thematic analysis was used to analyze the relevant literature and articles published in the period from 2011 to 2024. Thematic analysis was chosen because of its ability to successfully systematize key categories and create an adequate theoretical framework. The results of the research revealed eight key categories for the detection of disinformation: harm level, source checking, linguistic, syntactic, psycho-linguistic, style, visual and social context categories. These categories offer a systematic approach to recognizing disinformation from different perspectives, and the research itself emphasizes the importance of collaboration between people and analysis software. The research represents a comprehensive theoretical framework that not only contributes to the academic debate, but also serves as a foundation for future educational materials and experimental research.

Keywords: disinformation; disinformation detection; fact-checking; categories for checking disinformation; media environment; thematic analysis

Citation: Boko, Antea. 2024. Disinformation Detection: Developing a Categorical Framework Through Thematic Analysis. *Journalism and Media* 5: 1914–1924. https://doi.org/10.3390/journalmedia5040116

Academic Editor: Andreu Casero-Ripollés

Received: 21 October 2024
Revised: 5 December 2024
Accepted: 17 December 2024
Published: 20 December 2024

1. Introduction

The Internet, social media and smartphones have greatly transformed the way users communicate today. A change in the communication paradigm has enabled users to communicate with ease at any time regardless of geographic location. In addition, users today are faced with an enormous amount of information through the various channels they use. According to the Reuters Institute Digital News Report (Newman et al. 2022), for the most people, it is smartphones that are the first way they access news in the morning, and online platforms are the primary sources of information. Although this shows that information is more accessible to users than ever before, "some of this communication consists of false, inaccurate, and untrue information" (Søe 2017, p. 309). Hence, media audiences now live in a time of information disorder (Wardle and Derakhshan 2017; Wardle 2018) but also in the time of an infodemic (Rúas Araujo et al. 2022). Because of changes in the media, political and digital environment, many authors today use the term post-truth society to describe these issues (Carlson 2018). Consequently, 56% of internet users point out concerns when it comes to recognizing the accuracy of news in the online environment. Concern is manifested in those users who predominantly use social media as a source of information (Newman et al. 2023). In addition to the current challenges related to disinformation, the development of artificial intelligence, which creates such a realistic image that it becomes increasingly difficult to recognize that it is generated content, is certainly increasing the problem.

With the highlighted changes and problems, journalists are also in a situation where their verification skills must meet a higher level. Considering that disinformation is circulating rapidly on digital platforms, the role of fact-checking organizations and sites has been growing in recent years. Consequently, the fact-checking process has become a recent topic of interest to scientists in recent years. Their research includes different approaches, research methods and topics for detecting disinformation (e.g., Castillo et al. 2011; Cotter et al. 2022; Diaz Ruiz and Nilsson 2023; Grbeša Zenzerović and Nenadić 2022; Horne and Adali 2017; Kapantai et al. 2021; Pathak et al. 2021; Rubin et al. 2015; Rubin 2019; Shu et al. 2017a; Søe 2017; Tompkins 2019; Tsfati et al. 2020; Wardle and Derakhshan 2017; Zhang and Ghorbani 2020).

With the growth of interest in the topic of disinformation, numerous authors have described their definitions of this term. According to Diaz Ruiz and Nilsson (2023, p. 29), disinformation is "an adversarial campaign that weaponizes multiple rhetorical strategies and forms of knowing—including not only falsehoods but also truths, half-truths, and value-laden judgments—to exploit and amplify identity-driven controversies". Disinformation is the intentional presentation of misleading, false, fabricated and inaccurate content with the aim of misleading the audience and cause harm (Grbeša Zenzerović and Nenadić 2022; Pathak et al. 2021; Søe 2017; Wardle 2018). Based on existing definitions, Grbeša Zenzerović and Nenadić (2022, p. 11) identify "verifiability, intention and harm" as the key determinants of disinformation.

Summarizing the challenges in the media environment and the fact that disinformation is all around us, the goal of this study is to detect and describe categorical framework for disinformation detection. In this paper, the focus is on the systematization of key categories for checking disinformation, which enable a better and more detailed understanding of this phenomenon and provide the basis for the development of effective strategies for identifying false information. Precisely because of the stated focus and purpose, the method of thematic analysis was used. In a further part of the paper, the research method will be explained, and then the results through the key categories detected by this research will be discussed. Such a framework will be presented given that research shows that audiences are more likely to recognize disinformation if they have more data to verify (Paskin 2018). Also, the reason for conducting this research lies in the need for a comprehensive theoretical framework that can serve as a basis for further analysis and practical activities in the fight against disinformation. At the end of the paper, the presented categories form a framework that contributes to the academic understanding of the phenomenon, but also provides the basis for practical application in public education and further research.

2. Materials and Methods

As previously pointed out, this research aims to detect, describe and synthesize key categories for disinformation verification. Given that a significant problem of disinformation has been observed in the modern media environment, it is important to recognize a categorical framework through this paper that will help mitigate this challenge. This research aims to answer the research question—which are the key categories for checking disinformation and what are their characteristics and patterns?

For the purposes of this research, the method of thematic analysis will be used. This method is a qualitative research method that is used for "systematically identifying, organizing, and offering insight into patterns of meaning (themes) across a data set" (Braun and Clarke 2012, p. 57). So, in the thematic analysis, a specific text forms a set of data, and the codes are created in a way that the researcher develops them (Neuendorf 2018). Thematic analysis was chosen because it allows the researcher to synthesize and structure complex data to identify patterns and connections between them. Unlike content analysis, which is more focused on quantifying and classifying elements in text according to predefined categories, thematic analysis provides a more flexible framework for exploring meaning and context within the data. This is particularly useful in research such as this, where the goal is not only to identify, but also to understand the complex dimensions of disinformation across different categories. Thematic analysis allows for the development of new categories based on a thorough review of the data, making it an ideal method for research that requires interpretive understanding, such as disinformation detection.

To collect relevant data for the analysis of this topic, a search of different databases was first made. More specifically, the Scopus and Google Scholar databases were included in this research. The database search included articles from various sciences such as social science, natural science, and formal science. This shows that the topic of disinformation is present in wide spheres. Since the focus of the research is on the social and communicative aspects of disinformation, the inclusion of articles from other scientific disciplines could broaden the scope in a way that would reduce the specificity and depth of analysis within the chosen field. The exclusion of articles from technical, natural or other scientific fields enabled the analysis to be focused on thematic and theoretical issues relevant to the social sciences. So, for the purposes of this research, articles from social sciences, i.e., communication sciences, were included. To create a research sample, databases were searched using the key words "disinformation", "fact-checking" and "disinformation detection". The study selection criteria for this thematic analysis were based on several key factors that ensured the quality of the research. Studies were selected based on the relevance to the topic of disinformation and strategies for identifying it. Also, the included studies used valid methodological approaches, including qualitative and quantitative methods. They considered specific thematic categories and provided key insights into the analyzed categories. Studies that are recognized as significant in the academic community, i.e., frequently cited and recognized works within the disinformation research framework, were also included. The sample included works published in the period from 2011 to 2024. For the purposes of this research, the relevant literature was used, and therefore the emphasis is on key works published in the period from 2017.

After collecting the key literature, it was important to conduct a thematic analysis. The first step involved identifying data of potential interest and then systematically analyzing and segmenting them into codes. The coding process was performed manually. This was followed by searching for themes and identifying new patterns that resulted in a categorical framework for checking disinformation. Through the thematic analysis of the existing literature, the results of the research were detected, which will be presented in the next part of this paper.

3. Categorical Framework for Disinformation Detection

3.1. Harm Level Category

Many authors also tried to define different typologies of disinformation, misinformation and other false information (e.g., Kapantai et al. 2021; Kumar and Shah 2018; Lemieux and Smith 2018; Pamment et al. 2018; Pathak et al. 2021; Rubin et al. 2015; Rubin 2019; Tambini 2017; Tandoc et al. 2018; Wardle and Derakhshan 2017; Wardle 2018; Zannettou et al. 2019).

Wardle and Derakhshan (2017) and Wardle (2018) explained seven types of information disorder, and understanding of these categories significantly contributes to the detection of disinformation. All these types are based on level of harm, which is one of the key categories for understanding and checking disinformation. Although not intended to be harmful, (1) satire or parody still has the potential to mislead the audience. Furthermore, (2) false connection refers to the mismatch of the headlines, visuals and captions with the article (e.g., clickbait). A more harmful type of information disorder is (3) misleading content which is based on framing an issue or individual and (4) false context which is based on publishing accurate content out of the real context and thus misleads a person. Even more worrying forms of information disorder are (5) imposter content that misrepresents the original sources and (6) manipulated content that uses information or visuals to mislead the audience. The most harmful one is (7) fabricated content. It is completely new content (e.g., text, photo, video, web page) created with intention to deceive (Wardle and Derakhshan 2017; Wardle 2018). Based on the systematic literature review, Kapantai et al. (2021, p. 1317) gave a more detailed list of disinformation types, including clickbait, conspiracy theories, fabrication, misleading connection, hoax, bias or one-sidedness, imposter, pseudoscience, rumors, fake reviews and trolling.

3.2. Source Checking Category

As previously pointed out, in the fight against false information, many scientists are trying to provide an academic framework for identifying disinformation. After explaining different typologies of disinformation based on the level of harm, the next fact-checking category which will be presented in this paper is source. Given that digitalization has facilitated the publication of news in various decontextualized forms, today it is difficult to recognize whether it is a credible source of information (Rubin 2019). Certain sources of news and information in general have gained their reputation over time and are valued as reliable (Rubin et al. 2015).

The method of recognizing disinformation based on the credibility of the source implies the quality and believability (Zhou and Zafarani 2020). Also, according to Zhang and Ghorbani (2020), more authors believe that news sources are crucial in the process of detecting disinformation. Therefore, they suggest that it is important to check a domain and the URL, but also sections such as "About Us" or "Disclaimer", the date the news was published, other sources and their credibility, supporting resources such as statistical data, documents, external links and references (Zhang and Ghorbani 2020). Sitaula et al. (2020) suggest that articles without an author are more likely to be fake. Furthermore, the results show that authors who are labeled as credible will not cooperate with authors who are associated with writing disinformation. Author affiliation with recognized organizations is also a credibility detector (Sitaula et al. 2020). Research examining the audience's perception of disinformation detection has also shown the importance of verifying the source of the information itself (Acomi et al. 2021; Kyriakidou et al. 2022). Also, Hameleers et al. (2022) explain that users who utilize various sources of information to obtain news have an increased awareness of disinformation.

3.3. Linguistic Category

The process of recognizing disinformation is also based on language features, especially linguistic and syntactic features, which are a relevant and quality form for analyzing any form of false information (Horne and Adali 2017; Lebernegg et al. 2024; Shu et al. 2017a; Tompkins 2019; Zhang and Ghorbani 2020). These features "refer to the fundamental component, structure and semantics for natural language" (Zhang and Ghorbani 2020, p. 16). Tompkins (2019) explains lexical-based features as character-level and word-level features. To use this category in the identification of disinformation, it is necessary to pay attention to the word count, number of words per sentence, number of nouns, proper nouns, personal pronouns, possessive pronouns, determinants, cardinal numbers, adverbs, interjections, verbs, verb tenses, quantifying words, comparison words, exclamation marks, negations (e.g., no, never, not), swear words, online slang terms (e.g., lol, brb), interrogatives (e.g., how, what, why), stop words (e.g., the, is, on), punctuation, quotes, verb phrases and others (Horne and Adali 2017). Furthermore, the use of hashtags, bold words, question marks and exclamation points, reposting and emoticons also play an important lexical role on social media (Castillo et al. 2011; Zhang and Ghorbani 2020). Referring to the length of certain information, it is sometimes difficult to obtain a comparison between a short paragraph posted on a social network or a longer article (Rubin et al. 2015). Lexical-based features for identifying disinformation also include word readability and type–token ratio (Horne and Adali 2017; Zhang and Ghorbani 2020).

3.4. Syntactic Category

On the other hand, syntactic-based features are also called sentence-level features (Tompkins 2019; Zhang and Ghorbani 2020). In the literature, these features include the average sentence and post length, average sentence polarity, how punctuation is used, parts of speech tagging (POS) and other elements (Castillo et al. 2011; Horne and Adali 2017;

Shu et al. 2017a; Tompkins 2019; Zhang and Ghorbani 2020). Moreover, authors detected the domain-specific features that "include ratios of quoted words and external links, and the number of paragraphs and their average length in a document" (Potthast et al. 2017).

To check and analyze syntactic-based features, the neutral language processing (NLP) and machine learning (ML) programs are used. Stanford Parser software is used to recognize the complexity of a sentence, which calculates the depth of the syntax tree for each sentence (Horne and Adali 2017; Sabeeh et al. 2019; Shrestha and Spezzano 2021; Wang and Xu 2021). Furthermore, Python Natural Language Toolkit POS tagger is used for testing different characteristics in syntax (Horne and Adali 2017; Shrestha and Spezzano 2021; Verma et al. 2019), as well as probabilistic context free grammars (PCFG) for deep syntax experiment and analysis (Ashraf et al. 2021; Shu et al. 2017a).

3.5. Psycho-Linguistic Category

According to Tompkins (2019), psycho-linguistic features are used when verifying information. More specifically, the disinformation writing style is precisely characterized by the use of strong, sensationalist and clickbait expressions that encourage emotional engagement but also positive or negative sentiment (Tompkins 2019). Disinformation is used in a language and style that evokes various deep emotions in the audience, such as fear, sadness, anger, anxiety, empathy and curiosity (Savolainen 2023; Zhang and Ghorbani 2020) because of which it spreads more easily and rapidly. However, it should be considered that sometimes the titles of articles containing disinformation are "masked" and the use of clickbait is avoided (Pathak et al. 2021).

Furthermore, sentiment analysis is also used to check disinformation. It is a method used to identify and check patterns in different topics, actors or events that appear in the media (Barić et al. 2023). Considering sentiment analysis, authors define four categories—positive, negative, neutral and irrelevant (Feldvari et al. 2022). Also, words that indicate sentiment in the fact-checking process encompass the following: analytic, insightful, casual, discrepancy, tentative, certainty, differentiation, affiliation, power, reward, risk, personal concern, emotional tone and emotion words (Zhang and Ghorbani 2020).

Although it is untrue and false content, due to the use of the mentioned words and expressions, it is easier to attract people's attention (Humprecht et al. 2020). Moreover, it is a reason why a certain part of the audience often accepts false content. Such linguistic constructions precisely fulfill their purpose. They enable the recipients of the message to connect with the content or ideas on an emotional and mental level (Aguila Sánchez and Pereyra-Zamora 2022). In order to avoid certain misunderstanding, readers should read not only the headline of the article, but also the entire text because with disinformation there are often discrepancies between the headline and the text of the article (Zhang and Ghorbani 2020). To provide quality sentiment analysis, authors use the tool SentiStrength, which helps to detect positive or negative sentiment intensity of the text (Choraś et al. 2021; Horne and Adali 2017). Furthermore, for detailed psycho-linguistic features, authors used Linguistic Inquiry and Word Count (LIWC) (Ashraf et al. 2021; Shu et al. 2017a; Volkova and Jang 2018).

3.6. Style Category

One of the categories for identifying disinformation is also the style-based category which tends to find characteristic patterns of writing styles of disinformation authors (Potthast et al. 2017; Zafarani et al. 2019; Zhang and Ghorbani 2020; Zhou et al. 2019; Zhou and Zafarani 2020). Authors of disinformation try to create their content in a way that corresponds to the standard form of correct writing and thus mislead the reader. However, there are still ways to identify disinformation through writing style. In their study, Potthast et al. (2017) recognized patterns that indicate a characteristic style of disinformation writing, mostly conditioned by hatred, political predisposition and lack of interest. Although they concluded that style features alone are not sufficient to detect disinformation, they proved with their research that hyper-partisan content of both the left and the right is characterized by an extremist style of writing. In addition, they also recognized that satire is significantly different from the standard style, mostly due to the use of humor (Potthast et al. 2017).

Research by Zhou et al. (2019) show that the disinformation style is dominated by more informal language, subjectivity, emotional expressions, longer sentences, shorter words and unique verbs. News content is also an important detector of disinformation style (Zhou and Zafarani 2020). In order to recognize the characteristics of style, NLP and ML technologies are most often used (Zafarani et al. 2019).

3.7. Visual Category

After explaining the textual categories used to check disinformation, it is important to describe the visual categories as well. Namely, visual-based categories are often used as a tool for spreading various disinformation ideology. As with textual variants, visual disinformation is used with the aim of evoking a certain emotion in the recipient of the message and thus spreading propaganda (Shu et al. 2017a). Moreover, visual disinformation can be more dangerous than textual disinformation because it completely imitates reality and distorts human perception (Weikmann and Lecheler 2022). In their research, Weikmann and Lecheler (2022) described categories of visual disinformation in terms of manipulative sophistication and modal richness. Low sophistication is manifested in still images through actions of elimination and cropping. On the other hand, with moving images, changing the speed and adding video filters are used. Both still and moving images are often decontextualized and minimally edited, which is why they are called "cheap fake". High sophistication category for still images implies various photoshopping or doctoring actions, as well as misleading data visualizations. As for moving images, it is important to highlight the virtual performance, which implies different generations of video, sound, voice and so on. Finally, both still and moving images can be used as deepfakes (Weikmann and Lecheler 2022), especially when it comes to celebrities (Cao et al. 2020).

Forensic features are used to verify visual content. Firstly, such features are manifested through manipulation detection, which refers to the recognition of patterns created by copying, removing, moving and combining parts of visual content. Secondly, they include generation detection features that are used to recognize generated photos and video content (Cao et al. 2020). More specifically, authors point out that deepfake content can be detected by deep learning-based techniques, classical machine learning-based methods, statistical techniques and blockchain-based techniques (Rana et al. 2022). Deepfakes can be recognized by the absence of naturalistic eye blinking (Li et al. 2018) but also by blurred faces or with special effect, unnatural voice, lack of emotions, constant face swapping, irrelevant objects in background, unnatural behavior, lack of facial expression and abnormal mouth movements (Thaw et al. 2020). Goh (2024) also lists three categories that enable deepfakes to be recognized—the use of surface video and audio cues, the processing of the messages conveyed in the video, and the searching of external sources.

Finally, forensic features are based on re-compression detection, which means that fake visual content is subjected to double compression within manipulation of the content before saving and then during multiple downloads on social networks and when re-posting it. Also, visual disinformation detection is based on statistical characteristics such as number of images, popularity of sharing the content on social media, type of resolution and style, as well as on metadata (e.g., information about size and production) and external knowledge (e.g., timespan and platform credibility). However, using metadata is not always a useful way of checking because often textual data about visual content are not available (Cao et al. 2020).

3.8. Social Context Category

After explaining language-based categories, in the next few sections, social context features will be discussed. The first of them is called the network category, which is based on analyzing different groups that have emerged on social media platforms. These network groups are created by users with similar characteristics and they are connected based on, for example, their own interests, perspectives, topics, location, relationships, education and habits. Consequently, users who post similar content connect into characteristic networks (Shu et al. 2017b; Zhang and Ghorbani 2020). Based on the literature review by Shu et al. (2017a), there are stance networks, co-occurrence networks, friendship networks and diffusion networks that use different approaches of detection, for example, the clustering coefficient, SVD and network propagation algorithms.

Furthermore, authors describe distribution-based features that can help in finding ways of spreading disinformation. Building a propagation tree helps in this process (Castillo et al. 2011) as well as determining the "number of retweets/reposts for the original tweet/post, the fraction of tweets/posts that are retweeted for an online account, the in-degree/out-degree of an online user's ego net" (Zhang and Ghorbani 2020, p. 18). Also, it is important to pay attention to the features of different posts on social media, given that users are active in expressing their views on disinformation on these platforms. These post-based features include engagement level, group dynamics and temporal context (Shu et al. 2017a), but also the length of a message, the number of positive and negative sentiment words in a message, hashtags and reposts (Castillo et al. 2011). Temporal-based features are used to analyze online news publishing behavior. Such characteristics are useful for detecting suspicious activity and flagging disinformation. Common time features include the interval between posts, frequency of posting, replying, and commenting, time of posting and day of the week (Zhang and Ghorbani 2020).

Finally, the creator or used-based features will be tackled. Although content is often created by human users, it is recognized that disinformation is more often created and spread by non-human accounts, such as social bots and cyborgs that have intention to harm (Shu et al. 2017a; Zhang and Ghorbani 2020; Zhou and Zafarani 2020). However, the ideas and the content creation come from real people. These users therefore aim to manipulate society by spreading disinformation. In this category, it is important to pay attention to the general characteristics of the user, such as personality, location, verified registration information, description and so on (Shu et al. 2019; Zhang and Ghorbani 2020). Also, user-based features include the user's behavior and the credibility of the online account (Zhang and Ghorbani 2020), the number of followers, registration age and number of posts (Castillo et al. 2011). Hartwig et al. (2024) describe topical, formal and rhetorical characteristics in this category. On the other hand, sometimes disinformation is spread by people who are not aware that they are enacting this, or more precisely, they did not detect that it was disinformation so that must also be taken into consideration (Zhou and Zafarani 2020).

4. Discussion and Conclusions

Through this research, it was important to provide insight into a deeper understanding of the process of disinformation detection in the digital age. More specifically, it was important to determine the categories that enable the detection of disinformation using thematic analysis. Through the conducted research, patterns were systematized and then described through categories of harm level, source checking, linguistic, syntactic, psycholinguistic, style, visual and social context.

The first category detected by this research was the harm level category. This categorization aligns with Wardle and Derakhshan's (2017) seven types of information disorder, emphasizing that the higher the harm level, the greater the societal impact of the disinformation. This category is important because it indicates the impact and level of damage that disinformation can produce. The source-checking category is the closest to journalistic verification of information. Often, information, especially written as a journalistic article, appears very credible, but tends to mislead and manipulate. This is exactly why the authors point out that there are different ways to check the source (e.g., Zhang and Ghorbani 2020).

Recognizing disinformation is also revealed through linguistic and syntactic categories, because it is words and meanings that can reveal that something is false and misleading (Tompkins 2019; Horne and Adali 2017). In these categories, the use of different software programs based on artificial intelligence is essential. The psycho-linguistic category is based on sentiment analysis and is often easy to apply to texts involving charged emotions and sensationalism (e.g., Savolainen 2023). This category indicates the importance of education in media literacy and critical thinking.

The complexity of detecting disinformation is particularly prominent when it comes to the stylistic category. Namely, through the complete imitation of legitimate content using a specific writing style, it is very difficult to detect that there is fake content (Zhou et al. 2019). Nevertheless, a deeper understanding of this category enables the development of thinking and the ability to detect disinformation. As research shows, visual disinformation presents unique challenges, given its ability to distort perception through sophisticated manipulative techniques (Weikmann and Lecheler 2022). Precisely because of the realistic display created using artificial intelligence, it is important to constantly work on increasing people's ability to recognize artificially generated photos and videos, as well as tools that will recognize that it is disinformation. Through the category of social context, it is important to recognize the interconnected nature of online communities and the role of social media in spreading misinformation. Today, social media is the source of the largest amount of misinformation, so it is crucial to recognize that it comprises, for example, different bots and algorithms (e.g., Hartwig et al. 2024; Zhou and Zafarani 2020). Looking at this category, but also in general for all mentioned categories, there is a need for the responsible coordination of platforms to prevent the creation and spread of fake content in the digital media environment.

Through this study, a comprehensive framework for disinformation detection is presented, which breaks down the detection process into several categories. Each category deals with a certain problem, but also explains different ways to check the problems. Although it is important to emphasize that the findings indicate the need to use technological tools and artificial intelligence in disinformation detection, it is certainly important to have human supervision in the mentioned process. The public must be educated to successfully combat disinformation. Considering that disinformation is not only a one-dimensional problem, but includes textual, visual and social elements, it is necessary to constantly work on the development of categorical apparatus that can be used to check false content.

Some authors in similar research deal with several thematic categories in the same framework, which creates a synergy between them, and this significantly contributes to the strengthening of the entire analysis. For example, the category of source verification may be closely related to linguistic categories, as analysis of language and style often helps in assessing the credibility of information sources. Likewise, visual elements can be linked to social context, as images and videos often become viral content within certain social networks and online communities, further influencing their ability to spread disinformation. These relationships between categories allow deeper understanding of the disinformation detection process because they are not viewed in isolation but as interconnected dimensions of a complex problem.

In conclusion, although significant progress has been made in understanding and detecting disinformation, combating it requires ongoing collaboration between academics, technology platforms, governments and the public. Increasing media literacy and improving access to reliable sources of information are key to reducing the impact of disinformation in society. This study has a scientific contribution because it processes the relevant literature and creates a special theoretical and categorical framework in the field. Also, this research can be the basis for future educational material that will help the public to independently recognize disinformation. Future research could certainly include these categories in an experimental study. A more detailed investigation of these relationships between categories is recommended, as this could enable the development of new tools and methods for identifying disinformation based on the synergy between different thematic dimensions. Through additional studies that include experimental work, it is possible to develop a more comprehensive approach to disinformation analysis that considers their interconnectedness. In addition, this research aims to open opportunities for future research that will compare the process performed by fact-checking organizations on the one hand and journalists and media on the other.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The author declares no conflict of interest.

References

Acomi, Nicoleta, Luis Ochoa Siguencia, and Ovidiu Acomi. 2021. An Appropriate Set of Skills for Limiting the Spread of Fake News. *Revista Romaneasca Pentru Educatie Multidimensionala* 13: 71–80. [CrossRef]

Aguila Sánchez, Julio C., and Pamela Pereyra-Zamora. 2022. Infodemics in Mexico: A look at the Animal Político and Verificado fact-checking platforms. *Health Education Journal* 81: 982–92. [CrossRef]

Ashraf, Norman, Sabur Butt, Grigori Sidorov, and Alexander Gelbukh. 2021. CIC at CheckThat! 2021: Fake News detection Using Machine Learning And Data Augmentation. Paper presented at the CLEF 2021—Conference and Labs of the Evaluation Forum, Bucharest, Romania, September 21–24.

Barić, Ana, Laura Majer, David Dukić, Marijana Grbeša-Zenzerović, and Jan Šnajder. 2023. Target Two Birds with One STONE: Entity-Level Sentiment and Tone Analysis in Croatian News Headlines. Paper presented at the 9th Workshop on Slavic Natural Language Processing 2023 (SlavicNLP 2023), Dubrovnik, Croatia, May 6; pp. 78–85.

Braun, Virginia, and Victoria Clarke. 2012. Thematic Analyisi. In *APA Handbook of Research Methods in Psuchology, Vol 2. Research Desings: Quantitative, Qualitative, Neuropsychological, and Biological*. Edited by Harris Cooper, Paul M. Camic, Debra L. Long, A. T. Panter, David Rindskopf and Kenneth J. Sher. Washington, DC: American Psychological Association, pp. 57–71.

Cao, Juan, Peng Qi, Qiang Sheng, Tianyun Yang, Junbo Guo, and Jintao Li. 2020. Exploring the Role of Visual Content in Fake News Detection. In *Disinformation, Misinformation, and Fake News in Social Media*. Edited by Kai Shu, Suhang Wang, Dongwon Lee and Huan Liu. Cham: Springer, pp. 141–61.

Carlson, Matt. 2018. The information politics of journalism in a post-truth age. *Journalism Studies* 19: 1879–88. [CrossRef]

Castillo, Carlos, Marcelo Mendoza, and Barbara Poblete. 2011. Information credibility on Twitter. Paper presented at the 20th International Conference on World Wide Web, Hyderabad India, March 28–April 1; pp. 675–84.

Choraś, Michał, Konstantinos Demestichas, Agata Giełczyk, Álvaro Herrero, Paweł Ksieniewicz, Konstantina Remoundou, Daniel Urda, and Michał Woźniak. 2021. Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study. *Applied Soft Computing* 101: 107050. [CrossRef]

Cotter, Kelley, Julia R. DeCook, and Shaheen Kanthawala. 2022. Fact-checking the crisis: COVID-19, infodemics, and the platformization of truth. *Social Media+ Society* 8: 20563051211069048. [CrossRef]

Diaz Ruiz, Carlos, and Tomas Nilsson. 2023. Disinformation and echo chambers: How disinformation circulates on social media through identity-driven controversies. *Journal of Public Policy & Marketing* 42: 18–35.

Feldvari, Kristina, Milijana Mićunović, and Marija Blagović. 2022. COVID-19 Vaccines Disinformation and Conspiracy Theories: An Infodemic on Social Networks and Media in Croatia. *Medijske Studije* 13: 84–111. [CrossRef]

Goh, Dion Hoe-Lian. 2024. "He looks very real": Media, knowledge, and search-based strategies for deepfake identification. *Journal of the Association for Information Science and Technology* 75: 643–54. [CrossRef]

Grbeša Zenzerović, Marijana, and Iva Nenadić. 2022. *Strengthening Resilience to Disinformation: The State of Affairs and Guidelines for Action. Independent Study*. Zagreb: Agency for Electronic Media.

Hameleers, Michael, Anna Brosius, and Claes H. de Vreese. 2022. Whom to trust? Media exposure patterns of citizens with perceptions of misinformation and disinformation related to the news media. *European Journal of Communication* 37: 237–68. [CrossRef]

Hartwig, Katrin, Stefka Schmid, Tom Biselli, Helene Pleil, and Christian Reuter. 2024. Misleading information in crises: Exploring content-specific indicators for misleading information on Twitter from a user perspective. *Behaviour & Information Technology*: 1–34.

Horne, Benjamin, and Sibel Adali. 2017. This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. Paper presented at the International AAAI Conference on Web and Social Media, Montreal, QC, Canada, May 15–18; pp. 759–66.

Humprecht, Edda, Frank Esser, and Peter Van Aelst. 2020. Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics* 25: 493–516. [CrossRef]

Kapantai, Eleni, Androniki Christopoulou, Christos Berberidis, and Vassilios Peristeras. 2021. A systematic literature review on disinformation: Toward a unified taxonomical framework. *New Media & Society* 23: 1301–26.

Kumar, Srijan, and Neil Shah. 2018. False information on web and social media: A survey. *arXiv* arXiv:1804.08559.

Kyriakidou, Maria, Marina Morani, Stephen Cushion, and Ceri Hughes. 2022. Audience understandings of disinformation: Navigating news media through a prism of pragmatic scepticism. *Journalism* 24: 2379–96. [CrossRef]

Lebernegg, Noëlle, Jakob-Moritz Eberl, Petro Tolochko, and Hajo Boomgaarden. 2024. Do you speak disinformation? Computational Detection of Deceptive News-Like Content Using Linguistic and Stylistic Features. *Digital Journalism*, 1–24. [CrossRef]

Lemieux, Victoria, and Tyler D. Smith. 2018. Leveraging archival theory to develop a taxonomy of online disinformation. Paper presented at 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, December 10–13.

Li, Yuezun, Ming-Ching Chang, and Siwei Lyu. 2018. In ictu oculi: Exposing ai created fake videos by detecting eye blinking. Paper presented at 2018 IEEE International Workshop on Information Forensics and Security (WIFS), Hong Kong, China, December 11–13.

Neuendorf, Kimberly A. 2018. Content analysis and thematic analysis. In *Advanced Research Methods for Applied Psychology*. Edited by Paula Brough. London: Routledge, pp. 211–23.

Newman, Nic, Richard Fletcher, Craig T. Robertson, Kristen Eddy, and Rasmus Kleis Nielsen. 2022. *Reuters Institute Digital News Report 2022*. Oxford: Reuters Institute for the Study of Journalism.

Newman, Nic, Richard Fletcher, Kristen Eddy, Craig T. Robertson, and Rasmus Kleis Nielsen. 2023. *Reuters Institute Digital News Report 2023*. Oxford: Reuters Institute for the Study of Journalism.

Pamment, James, Howard Nothhaft, and Alicia Fjällhed. 2018. *Countering Information Influence Activities: The State of the Art*. Lund: Department of Strategic Communication, Lund University.

Paskin, Danny. 2018. Real or fake news: Who knows? *The Journal of Social Media in Society* 7: 252–73.

Pathak, Archita, Rohini K. Srihari, and Nihit Natu. 2021. Disinformation: Analysis and identification. *Computational and Mathematical Organization Theory* 27: 357–75. [CrossRef] [PubMed]

Potthast, Martin, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein. 2017. A stylometric inquiry into hyperpartisan and fake news. *arXiv* arXiv:1702.05638.

Rana, Md Shohel, Mohammad Nur Nobi, Beddhu Murali, and Andrew H. Sung. 2022. Deepfake detection: A systematic literature review. *IEEE Access* 10: 25494–513. [CrossRef]

Rubin, Victoria L. 2019. Disinformation and misinformation triangle: A conceptual model for "fake news" epidemic, causal factors and interventions. *Journal of documentation* 75: 1013–34. [CrossRef]

Rubin, Victoria L., Yimin Chen, and Nadia K. Conroy. 2015. Deception detection for news: Three types of fakes. *Proceedings of the Association for Information Science and Technology* 52: 1–4. [CrossRef]

Rúas Araujo, José, John P. Wihbey, and Daniel Barredo-Ibáñez. 2022. Beyond Fake News and Fact Checking: A Special Issue to Understand the Political, Social and Technological Consequences of the Battle Against Misinformation and Disinformation. *Journalism and Media* 3: 254–56. [CrossRef]

Sabeeh, Vian, Mohammed Zohdy, and Rasha Al Bashaireh. 2019. Enhancing the fake news detection by applying effective feature selection based on semantic sources. Paper presented at 2019 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, December 5–7.

Savolainen, Reijo. 2023. Assessing the credibility of COVID-19 vaccine mis/disinformation in online discussion. *Journal of Information Science* 49: 1096–110. [CrossRef] [PubMed]

Shrestha, Anu, and Francesca Spezzano. 2021. Textual characteristics of news title and body to detect fake news: A reproducibility study. In *Advances in Information Retrieval: 43rd European Conference on IR Research, ECIR 2021, Virtual Event, March 28–April 1, 2021, Proceedings, Part II 43*. Cham: Springer International Publishing.

Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017a. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter* 13: 22–36. [CrossRef]

Shu, Khai, Suhang Wang, and Huan Liu. 2017b. Exploiting tri-relationship for fake news detection. *arXiv* arXiv:1712.07709.

Shu, Khai, Xinyi Zhou, Suhang Wang, Reza Zafarani, and Huan Liu. 2019. The role of user profiles for fake news detection. Paper presented at the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Vancouver, BC, Canada, August 27–30.

Sitaula, Niraj, Chilukuri K. Mohan, Jennifer Grygiel, Xinyi Zhou, and Reza Zafarani. 2020. Credibility-Based Fake News Detection. In *Disinformation, Misinformation, and Fake News in Social Media*. Edited by Kai Shu, Suhang Wang, Dongwon Lee and Huan Liu. Cham: Springer, pp. 163–82.

Søe, Sille Obelitz. 2017. Algorithmic detection of misinformation and disinformation: Gricean perspectives. *Journal of Documentation* 74: 309–32. [CrossRef]

Tambini, Damian. 2017. *Fake News: Public Policy Responses. Media Policy Brief 20*. London: Media Policy Project, London School of Economics and Political Science.

Tandoc, Edson C., Jr., Zheng Wei Lim, and Richard Ling. 2018. Defining "fake news" A typology of scholarly definitions. *Digital Journalism* 6: 137–53. [CrossRef]

Thaw, Nyein Nyein, Thin July, Aye Nu Wai, Dion Hoe-Lian Goh, and Alton Y. K. Chua. 2020. Is it real? A study on detecting deepfake videos. *Proceedings of the Association for Information Science and Technology* 57: e366. [CrossRef]

Tompkins, Jillian. 2019. Disinformation Detection: A review of linguistic feature selection and classification models in news veracity assessments. *arXiv* arXiv:1910.12073.

Tsfati, Yariv, Hajo G. Boomgaarden, Jesper Strömbäck, Rens Vliegenthart, Alyt Damstra, and Elina Lindgren. 2020. Causes and consequences of mainstream media dissemination of fake news: Literature review and synthesis. *Annals of the International Communication Association* 44: 157–73. [CrossRef]

Verma, Abhishek, Vanshika Mittal, and Suma Dawn. 2019. FIND: Fake information and news detections using deep learning. Paper presented at 2019 Twelfth International Conference on Contemporary Computing (IC3), Noida, India, August 8–10.

Volkova, Svitlana, and Jin Yea Jang. 2018. Misleading or falsification: Inferring deceptive strategies and types in online news and social media. Paper presented at Companion Proceedings of the The Web Conference 2018, Lyon, France, April 23–27; pp. 575–83.

Wang, Yingxu, and James Y. Xu. 2021. An autonomous semantic learning methodology for fake news recognition. Paper presented at 2021 IEEE International Conference on Autonomous Systems (ICAS), Montreal, QC, Canada, August 11–13.

Wardle, Claire. 2018. The need for smarter definitions and practical, timely empirical research on information disorder. *Digital journalism* 6: 951–63. [CrossRef]

Wardle, Claire, and Hossein Derakhshan. 2017. *Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking*. Strasbourg: Council of Europe.

Weikmann, Teresa, and Sophie Lecheler. 2022. Visual disinformation in a digital age: A literature synthesis and research agenda. *New Media & Society* 25: 3696–713.

Zafarani, Reza, Xinyi Zhou, Kai Shu, and Huan Liu. 2019. Fake news research: Theories, detection strategies, and open problems. Paper presented at the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, August 4–8.

Zannettou, Savvas, Michael Sirivianos, Jeremy Blackburn, and Nicolas Kourtellis. 2019. The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality (JDIQ)* 11: 1–37. [CrossRef]

Zhang, Xichen, and Ali A. Ghorbani. 2020. An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management* 57: 102025.

Zhou, Xinyi, and Reza Zafarani. 2020. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)* 53: 1–40. [CrossRef]

Zhou, Xinyi, Atishay Jain, Vir V. Phoha, and Reza Zafarani. 2019. Fake news early detection: An interdisciplinary study. *arXiv* arXiv:1904.11679.