Key Characteristics of Health Misinformation and Disinformation on Social Media and Automated Machine Learning Techniques for Detection

Lakshika J. Paiva, Faculty of Engineering, University of Ottawa

Kimberley Nault (Prof), Telfer School of Management, University of Ottawa

1.0 INTRODUCTION

1.1 Background and Problem

In recent years, social media networks such as Facebook, Twitter, YouTube, and Instagram have appeared as new ways to access health information. Consequently, they offer convenient access to online content, facilitating the rapid dissemination of health-related misinformation and disinformation within global communities, especially during pandemics, health crises, and infectious disease outbreaks. This becomes dangerous because people heavily rely on social media for health information (Zhang et al., 2021) and they are unlikely to fact-check with a health professional (Neely et al., 2021). But "the quality of this kind of information is not guaranteed" (Madathil et al., 2015; Sayakhot & Carolan-Olah, 2016) because anyone can directly post and share health information (Zhu, 2016). The consequences are severe as health misinformation has the potential to cause serious harm to public health, from vaccine hesitancy to adoption of unproven disease treatments, decreased trust in public health systems (Nascimento et al., 2022), and increased hate speech towards ethnic groups or medical experts (Schlicht et al., 2023).

Before delving further, it is important to understand that there are two distinct interpretations of health misinformation and disinformation in the literature (Tandoc et al., 2018; Guo, 2020; Nascimento et al., 2022). However, in most studies, the umbrella term 'health misinformation' is used to signify health information on social media that is false, inaccurate, or misleading (Nascimento et al., 2022; Zhang et al., 2021; Zhong, 2023). Hence, in this study we use the term 'health misinformation' and its variations appearing in research to represent both misinformation and disinformation.

1.2 The Need

Given the serious consequences of health misinformation on social media, conducting an in-depth understanding of its characteristics in content and dissemination is the first step toward mitigating the surge of health misinformation (Zhong, 2023). Although most existing studies focuses on the spread and influence of health misinformation, rarely effort has been conducted to identifying the characteristics (Liu et al., 2019). Further, little is known about the key characteristics of health misinformation on social media itself (Zhang et al., 2021). Hence, this research focuses on exploring the main characteristics of health misinformation on social media.

A plethora of tools and techniques have been introduced to combat health misinformation on social media including fact-checking and automated detection mechanisms. However, fact-checking is a laborious process (Chen et al., 2023) and besides, amateurs do not have sufficient health literacy to identify accurate health information. As a result, "there is a need to develop automated solutions that can assist both experts and non-experts in discerning between genuine and non-genuine health information" (Di Sotto & Viviani, 2022). Literature presents large number of automated detection techniques using Machine Learning (ML) however, despite considerable work devoted to suppressing health misinformation on social media, it has increased in recent years (Zhong, 2023). Further, it is essential to analyse the characteristics of health-related misinformation on social media and to design automated detection tools (Liu et al., 2019). Therefore, this study explores automated detection techniques of health misinformation.

Accordingly, following research questions (RQ) are presented for this literature review:

- **RQ1:** What are the main characteristics of health misinformation spreading in social media?
- **RQ2:** What are the automated detection techniques of health misinformation on social media?

1.3 Significance of this Research

This study is significant because it is designed to help characterise and automatically detect health misinformation on social media as the first step to protecting social media users and vulnerable groups from causing serious harm to their health. Moreover, the study contributes to close research gaps in discovering misinformation features and improving automatic detection techniques through impactful feature recognition, so that it can improve detection. As a result, the outcomes are expected to guide the work of social media giants and healthcare technology enterprises in reducing exposure to and mitigating chances of misinformation consumption. Further, implications benefit and ease the life of fact-checkers and experts involved in verifying large amounts of real health information.

2.0 METHODS

The primary databases utilized in this literature search were Scopus, Web of Science, IEEE, ACM and ScienceDirect. A few relevant articles were identified through Google Scholar search. The search strategy was designed using Boolean queries including relevant terms like ‘misinformation’, ‘disinformation’, ‘fake news’, ‘characteristics’, ‘features’, ‘detect*’, ‘identif*’ ‘tools’, ‘techniques’, ‘technolog*’, ‘machine learning’, ‘social media’, ‘Facebook’, ‘Twitter’, ‘YouTube’, ‘Instagram’, used both separately and in combination, along with their variations as appearing in publications. In certain instances, the results were filtered using journal or conference names to improve relevance of results to internet or medical research. Moreover, the seminal articles corresponding to each RQ were used as seed papers in the ‘Inciteful’ tool (Weishuhn, 2023) to uncover similar papers. The results were screened to ensure they were relevant peer-reviewed journals or conferences based on empirical research and published between 2019 and 2023 to ensure quality.

3.0 LITERATURE REVIEW

This section critically evaluates the literature according to the research questions on characteristics and automated detection techniques and presents the evaluation in a concept-centric manner.

3.1 RQ1: Main characteristics of health misinformation spreading on social media

The literature identifies a broad range of characteristics (i.e., features) of health misinformation classified according to message characteristics, propagation pattern, and information source, structure of social media, temporal features, and receiver. However, based on the frequency, message features, propagation pattern and information source stood out as the primary features that appeared in most of the studies with more granularity. Due to time and space constraints, this literature review focuses on these three commonly appearing characteristics as presented in Table 1.

Table 1: Concept Matrix for Characteristics

Articles	Concepts				
	Message characteristics/ features			**Propagation pattern**	**Information source**
	Linguistic/ Writing Strategy	Sentiments/ Emotions	Content quality	Engagement / Virality	-

Zhang et al. (2021)		✓			✓
Zhong (2023)		✓	✓	✓	✓
Zhou et al. (2021)	✓	✓			
Li et al. (2022)			✓		✓
Chen et al. (2023)			✓		✓
Ngai et al. (2022)	✓		✓	✓	
Mahlous & Al-Laith (2021)		✓			
Zhao et al. (2022)		✓	✓	✓	✓

3.1.1 Message characteristics

Much of the literature focuses on characterising misinformation on social media using message features. Message characteristics have four prominent sub features including linguistic or writing strategy, sentiments and emotions, and content quality.

3.1.1.1 Linguistic or Writing Strategy

Few of the literature focuses on the linguistic or writing strategy to characterise misinformation on social media. Primarily, linguistic characteristics in health misinformation are classified into word, sentence and content-level characteristics or strategies.

There are different linguistic or writing strategies followed by misinformation with varied intentions. At the word level, creators often use persuasive, comparative, and emotional words with the intention of providing social support and creating trust in the creator, which in turn nurtures individuals' intention to repost misinformation (Zhou et al., 2021). In comparison, conversational or personal tones are adopted to induce emotions like fear, anxiety, and mistrust among the public to make it more accessible (Ngai et al., 2022). Adding to the review, writing strategies also consist of mimicking the language features and format of mainstream news and scientific reports, using an informal conversational style, and employing amplification (Jamison et al., 2020) for credibility, for instance like capitalizing all letters in the first word of a sentence (Ngai et al., 2022).

While writing strategies used are varied, the primary intention of all the writing strategies is to facilitate widespread dissemination in social media. Notably, it is evident that linguistics or writing strategies are under researched in the literature of characteristics and needs more research. A significant highlight is that writing strategies are often used in combination with emotions, which is another significant feature discussed in the next section.

3.1.1.2 Sentiments or Emotions

Majority of the literature reported the use of sentiments or emotional words in misinformation or provoking different types of emotions, primarily as negative, positive, or neutral.

Emotions in misinformation play a critical role in deceiving social media users (Ghanem et al., 2020), primarily to enable widespread propagation. In support, three studies found that intense or strong emotions that trigger high arousal in the public enable extensive propagation (Guo et al., 2019; Vosoughi et al., 2018; Zhou, 2022). However, these studies did not specify which emotions, positive or negative, triggered the propagation.

More scrutiny reveals that compared to correct information, misinformation contains more negative emotion words (Long et al., 2017). Adding to this fact, the relationship between negative emotions and misinformation dissemination was found to be positive and significant (Zhou et al., 2021). Further evidence shows, "tweets related to fake news are more negative and have strong sentiment polarity in comparison with genuine news" (Mahlous & Al-Laith, 2021).

In contrast to the widespread view about negative emotions (Xu et al., 2020), one study claimed misinformation tended to express positive emotions to attract public attention (Zhao et al., 2022). However, there is a need for more strong evidence to justify the latter.

Overall, strong evidence on negative emotions or sentiments and dissemination of misinformation makes it clear that negative emotions are a key characteristic of misinformation on social media.

3.1.1.3 Content Quality

Much of the literature identifies content quality as a key characteristic of health misinformation.

The critical content quality feature of health misinformation is 'lack of credibility' (Li et al., 2022). According to a study, credibility was primarily based on the source of health information (Chen et al., 2023). Adding to this, other studies claim that higher the perceived credibility of misinformation is, the stronger the social media users' willingness to forward the message (Halpern et al., 2019; Kim et al., 2019; Seah & Weimann, 2020).

In addition, lack of accuracy plays a significant role in misinformation denoted by bad grammar, incomplete content and unfair opinions (Li et al., 2022). Other studies add value, claiming that 'content repetition' may contribute to the perceived accuracy of misinformation and the likelihood of its going viral on social media (Chen et al., 2023; Lee & Shin, 2021; Pennycook et al., 2018).

While the lack of credibility and accuracy were identified as a key determinant of content quality, the measures of credibility were poorly discussed in the studies. Commonly, both features pointed to creating a false perception of content quality that enabled dissemination of misinformation. Overall, the discussion of content quality took several other sub characteristics like reliability and support, however, they need more evidence to justify their representation in misinformation.

3.1.2 Propagation or Spreading Pattern

A few literature sources present that spreading patterns are a key characteristic of health misinformation. Propagation or spreading pattern is identified by the 'engagement' or 'virality' of health misinformation on social media (Ngai et al., 2022) such as likes, comments and shares.

There is widespread evidence to support the spreading pattern as a strong characteristic of misinformation. For instance, stronger sentiments produce larger amounts of discussion with evidence that found fake health news generated higher response compared to real news (Zhong, 2023). Similarly, previous studies have revealed that emotional texts facilitate individuals' social media engagement behaviors, including likes (Ji et al., 2019), shares (Tang et al., 2019) and comments (Ji et al., 2019). Adding to that, misinformation that imitated the format of news media and scientific reports were associated with likes (Ngai et al., 2022).

The literature shows strong evidence to justify that propagation pattern is a key characteristic of health misinformation. Importantly, it was evident that propagation did not occur alone, instead it was coupled with other characteristics such as emotions, sentiments and writing styles which are evaluated above. In combination with previous characteristics, there is a strong justification for propagation patterns as a key characteristic of health misinformation.

3.1.3 Information source

Several literature sources identified information sources as a key characteristic of misinformation.

A study investigating the relationship between source credibility and propensity to share misinformation on social media found a strong relationship (Nadarevic et al., 2020). Similarly, "most previous studies highlight the importance of source credibility or senders' trustworthiness as factors affecting social media users' belief in misinformation content (Kim et al., 2019; Nadarevic et al., 2020) and subsequent susceptibility to sharing misinformation posts (Chen et al., 2023).

Overall, information sources have a direct impact on misinformation propagation, given that their credibility and trustworthiness is a key determining factor to share misinformation on social media.

In summary of the section, it is notable that although writing strategies have gaps in literature, message characteristics, spreading patterns and information source are strong features of health misinformation. The characteristics often operate in combination with one another, notably the propagation patterns.

3.2 RQ2: Automated detection techniques of health misinformation on social media

Much of the literature on automated detection techniques use machine learning techniques for detecting health misinformation. In this review we focus on the four techniques that emerged from the research including text, emotion, and behavioural analysis (refer Table 2), because they represent characteristics of health misinformation and facilitates a better basis for comparison.

3.2.1 Text Analysis

3.2.1.1 Topics and Themes

Much of the literature related to health misinformation detection techniques leverages topics and themes (Li et al., 2022); Zhao et al., 2021; Choudrie et al., 2021; Gupta et al., 2022) which scrutinise contents at the word, phrase, or sentence levels.

Topics and themes analysis reveals that real and fake health information both share similarities and differences. Moreover, reliable health information covers a broad spectrum of objective health topics (Gupta et al., 2022), often linked to authoritative sources and backed by scientific evidence (Choudrie et al., 2021). In contrast, fake news tends to focus on narrower subjects (Liu et al., 2019) often rooted in opinions (Gupta et al., 2022) and experiences (Liu et al., 2019), lacking links to authoritative sources or scientific evidence (Choudrie et al., 2021). Further, analysis of topics and themes highlights that unreliable health articles are often motivated by profit generation, while reliable ones aim to promote healthy lifestyles and accurate disease treatment (Liu et al., 2019).

Given the success of the majority of the studies in using topics and themes to distinguish between genuine and non-genuine health information, it can be deduced that the techniques have proven essential in enabling the pinpointing of health misinformation.

3.2.1.2 Linguistic Features

Literature underscores the significance of linguistic features in identifying deceptive narratives by scrutinizing elements, such as message length, structure, punctuation, and grammar (Choudrie et al., 2021; Liu et al., 2019; Safarnejad et al., 2021; Zhao et al., 2021).

Health misinformation tends to be much longer on average (Zhao et al., 2021), however similarly reliable articles also tend to be longer because the latter demonstrates medical facts and references (Liu et al., 2019). Moreover, titles of misinformation tend to be longer and have more exclamation marks (Liu et al., 2019). However, two out of the three studies suggest that linguistic features may not be strong detectors because content styles are evolving (Choudrie et al., 2021) or they were less informative in detecting health misinformation (Zhao et al., 2021). Based on these points, it can be deduced that linguistic features are not strong factors for health misinformation detection.

Table 2: Concept Matrix for Machine Learning Techniques

Articles	**Concepts**				
	Text analysis		**Emotion analysis**		**Behavioral analysis**
	Topics & themes	Linguistic features	Emotions	Sentiments	-
Gupta et al. (2022)	✓		✓		
Zhao et al. (2021)	✓	✓	✓	✓	✓
Liu et al. (2019)	✓	✓			
Safarnejad et al (2021)		✓	✓		✓
Choudrie et al. (2021)	✓				

Ramakrishnan & Balakrishnan (2022)				✓	

3.2.2 Emotion and Sentiment Analysis

The application of emotions and sentiment analysis in health misinformation detection is gaining attention in the literature. Emotions, such as anger, fear, sadness, and joy, are analysed, and sentiment analysis categorizes text into positive, neutral, and negative sentiments, offering insights into people's thoughts and opinions (Ramakrishnan & Balakrishnan, 2022).

Studies show that negative emotions are more prevalent in fake news, which is a valuable feature for distinguishing misinformation. In a covid-19 based misinformation study, fear was the dominant negative emotion with the next common emotion being sadness (Gupta et al., 2022). Comparably, another study claims that both misinformation and legitimate information expressed similar emotions (Zhao et al., 2021), however they did not expose if the emotion was negative or positive like in the former study. Moreover, sentiment analysis is considered an effective tool for misinformation differentiation (Ramakrishnan & Balakrishnan, 2022). However, further research suggests that sentiment features have been seen as effective features in detecting online rumors and fake reviews (Rout et al., 2017), and it is also widely used in health misinformation detection. Overall, based on the outcome of these studies, we deduce that emotions and sentiments are strong machine learning techniques for health misinformation detection.

3.2.3 Behavioral Analysis or Network Analysis:

According to the literature, behavioral analysis techniques have gained prominence, focusing on network, relationship, and interaction-based machine learning tools, within the realm of social media. These techniques involve user-based features such as the number of threads, replies (Zhao et al., 2021), followers, friends, and verification status (Safarnejad et al., 2021).

A study claims that user behavioural features were more powerful and informative in in detection compared to linguistic and semantic features (Zhao et al., 2021), although it does not provide specific reasons. Comparatively, another study identified user accounts frequently disseminating misinformation, including credible ones like ‘The Economist’ due to a lack of understanding of evolving health issues (Safarnejad et al., 2021). However, there lacks strong evidence on the impact of user behaviour on misinformation dissemination.

While there's positive evidence supporting the role of user behavioral features in misinformation detection, there are significant research gaps, especially regarding the detailed analysis of elements like the number of threads and replies.

4.0 DISCUSSION AND CONCLUSIONS

4.1 RQ1: Main characteristics of health misinformation spreading on social media

The literature highlights message characteristics, propagation patterns, and information sources as key factors in health misinformation on social media. Emotions within message features strongly influence people, while writing strategies, an area with significant research gaps, play a pivotal role when combined with emotions in spreading misinformation. These characteristics often interact, with propagation patterns closely linked to writing strategy, content quality, emotions, and information source, enhancing dissemination. Automated health misinformation detection systems can utilize these factors for greater impact on curbing the spread.

4.2 RQ2: Automated detection techniques of health misinformation on social media

The literature on automated detection techniques offers valuable insights into automated detection. The integration of text analysis and sentiment analysis into machine learning frameworks appears to enhance overall accuracy in detecting misinformation, offering a potentially powerful tool in combating false health information. Moreover, the multi-faceted approach to health misinformation detection involving text analysis, emotion and sentiment analysis, and behavioral analysis is essential. The combination of these techniques within machine learning frameworks provides a more comprehensive and effective approach to safeguarding public health and ensuring the dissemination of accurate information in the digital age.

4.3 Combined implications

This study has significant implications for developing health misinformation detection systems on social media. Emotions and sentiments were found to be robust characteristics in the feature analysis (RQ1) as well in automated detection techniques (RQ2).

However, there is a notable disparity between other identified characteristics (RQ1) and their utilization in detection techniques (RQ2). While linguistic styles and writing strategies have limited research, detection systems predominantly rely on topics and themes (relying on writing strategies), indicating the need for more exploration. Content quality, although vital, is rarely used in current detection mechanisms. Propagation features and information sources align with user behavioral detection mechanisms but require further research for effective automated detection.

These key findings guide and impact the work of social media giants and healthcare technology enterprises in reducing exposure to and mitigating chances of misinformation consumption.

Additional implications present that, youth are more susceptible to health misinformation on social media than older adults as the latter group's reliance on classic communication channels made them relatively immune to online health misinformation.

4.4 Gaps and limitations

While some aspects like message, propagation, and sources have been explored, there is a need for more comprehensive investigations into nuanced features, such as temporal evolution, image analysis, and the role of information consumers. The limitations include the underrepresentation of platforms like Facebook, YouTube, and Instagram in the literature, making it challenging to generalize the features identified to all platforms.

4.5 Recommendations for future research

Future research should delve deeper into linguistic styles and writing strategies to inform the development of automated health misinformation detection systems. There's also an opportunity for more focused research on specific social media platforms like Facebook, YouTube, and Instagram. Additionally, studying the spread of health misinformation from one platform to another can be a valuable research area. Lastly, improvements in the communication of technical results in everyday language can enhance the effectiveness of research in detection techniques and advance the field.

4.6 Conclusion

In conclusion, the study identifies three key characteristics of health misinformation on social media and three key techniques used in the automated detection of health misinformation. In comparison of the findings of the two research questions, emotions and sentiments stand out as having a direct overlap and the strongest characteristic useful for health misinformation detection. It also presents that a combination of characteristics is required for automated detection systems to function with success as health misinformation characteristics do not appear in isolation, but often function together.

5.0 REFERENCES

Chen, S., Xiao, L., & Kumar, A. (2023). Spread of misinformation on social media: What contributes to it and how to combat it. *Computers in Human Behavior*, *141*, 107643. https://doi.org/10.1016/j.chb.2022.107643

Choudrie, J., Banerjee, S., Kotecha, K., Walambe, R., Karende, H., & Ameta, J. (2021). Machine learning techniques and older adults processing of online information and misinformation: A covid 19 study. *Computers in Human Behavior*, *119*, 106716. https://doi.org/10.1016/j.chb.2021.106716

Di Sotto, S., & Viviani, M. (2022). Health Misinformation Detection in the Social Web: An Overview and a Data Science Approach. *International Journal of Environmental Research and Public Health*, *19*(4), Article 4. https://doi.org/10.3390/ijerph19042173

Ghanem, B., Rosso, P., & Rangel, F. (2020). An Emotional Analysis of False Information in Social Media and News Articles. *ACM Transactions on Internet Technology*, *20*(2), 1–18. https://doi.org/10.1145/3381750

Guo, C., Cao, J., Zhang, X., Shu, K., & Yu, M. (2019). Exploiting Emotions for Fake News Detection on Social Media. *ArXiv*. https://www.semanticscholar.org/paper/Exploiting-Emotions-for-Fake-News-Detection-on-Guo-Cao/1817701ad21aaadcaed95434899a25c294767a01

Gupta, A., Li, H., Farnoush, A., & Jiang, W. (2022). Understanding patterns of COVID infodemic: A systematic and pragmatic approach to curb fake news. *Journal of Business Research*, *140*, 670–683. https://doi.org/10.1016/j.jbusres.2021.11.032

Halpern, D., Valenzuela, S., Katz, J., & Miranda, J. P. (2019). From Belief in Conspiracy Theories to Trust in Others: Which Factors Influence Exposure, Believing and Sharing Fake News. In G. Meiselwitz (Ed.), *Social Computing and Social Media. Design, Human Behavior and Analytics* (pp. 217–232). Springer International Publishing. https://doi.org/10.1007/978-3-030-21902-4_16

Jamison, A., Broniatowski, D. A., Smith, M. C., Parikh, K. S., Malik, A., Dredze, M., & Quinn, S. C. (2020). Adapting and Extending a Typology to Identify Vaccine Misinformation on Twitter. *American Journal of Public Health*, *110*(Suppl 3), S331–S339. https://doi.org/10.2105/AJPH.2020.305940

Ji, Y. G., Chen, Z. F., Tao, W., & Cathy Li, Z. (2019). Functional and emotional traits of corporate social media message strategies: Behavioral insights from S&P 500 Facebook data. *Public Relations Review*, *45*(1), 88–103. https://doi.org/10.1016/j.pubrev.2018.12.001

Kim, J.-H., Bock, G.-W., Sabherwal, R., & Kim, H.-M. (2019). Why Do People Spread Online Rumors? An Empirical Study. *Asia Pacific Journal of Information Systems*, *29*(4), 591–614.

Lee, E.-J., & Shin, S. Y. (2021). Mediated Misinformation: Questions Answered, More Questions to Ask. *American Behavioral Scientist*, *65*(2), 259–276. https://doi.org/10.1177/0002764219869403

Li, Y., Fan, Z., Yuan, X., & Zhang, X. (2022). Recognizing fake information through a developed feature scheme: A user study of health misinformation on social media in China. *Information Processing & Management*, *59*(1), 102769. https://doi.org/10.1016/j.ipm.2021.102769

Liu, Y., Yu, K., Wu, X., Qing, L., & Peng, Y. (2019). Analysis and Detection of Health-Related Misinformation on Chinese Social Media. *IEEE Access*, *7*, 154480–154489. https://doi.org/10.1109/ACCESS.2019.2946624

Long, Y., Lu, Q., Xiang, R., Li, M., & Huang, C.-R. (2017, November 1). *Fake News Detection Through Multi-Perspective Speaker Profiles*. International Joint Conference on Natural Language Processing. https://www.semanticscholar.org/paper/Fake-News-Detection-Through-Multi-Perspective-Long-Lu/4c2d04e3c977839fc524a6bda4d18bdefcc8ce2b

Madathil, K. C., Rivera-Rodriguez, A. J., Greenstein, J. S., & Gramopadhye, A. K. (2015). Healthcare information on YouTube: A systematic review. *Health Informatics Journal*, *21*(3), 173–194. https://doi.org/10.1177/1460458213512220

Mahlous, A. R., & Al-Laith, A. (2021). Fake News Detection in Arabic Tweets during the COVID-19 Pandemic. *International Journal of Advanced Computer Science and Applications*, *12*(6). https://doi.org/10.14569/IJACSA.2021.0120691

Nadarevic, L., Reber, R., Helmecke, A. J., & Köse, D. (2020). Perceived truth of statements and simulated social media postings: An experimental investigation of source credibility, repeated exposure, and presentation format. *Cognitive Research: Principles and Implications*, *5*(1), 56. https://doi.org/10.1186/s41235-020-00251-4

Neely, S., Eldredge, C., & Sanders, R. (2021). Health Information Seeking Behaviors on Social Media During the COVID-19 Pandemic Among American Social Networking Site Users: Survey Study. *Journal of Medical Internet Research*, *23*(6), e29802. https://doi.org/10.2196/29802

Ngai, C. S. B., Singh, R. G., & Yao, L. (2022). Impact of COVID-19 Vaccine Misinformation on Social Media Virality: Content Analysis of Message Themes and Writing Strategies. *Journal of Medical Internet Research*, *24*(7), e37806. https://doi.org/10.2196/37806

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology. General*, *147*(12), 1865–1880. https://doi.org/10.1037/xge0000465

Ramakrishnan, K., & Balakrishnan, V. (2022). SentiMage: A Sentiment-Image-based COVID-19 Health Misinformation Detection using Machine Learning. *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, 1–5. https://doi.org/10.1109/ICECCME55909.2022.9987818

Rout, J. K., Singh, S., Jena, S. K., & Bakshi, S. (2017). Deceptive review detection using labeled and unlabeled data. *Multimedia Tools and Applications*, *76*(3), 3187–3211. https://doi.org/10.1007/s11042-016-3819-y

Safarnejad, L., Xu, Q., Ge, Y., & Chen, S. (2021). A Multiple Feature Category Data Mining and Machine Learning Approach to Characterize and Detect Health Misinformation on Social Media. *IEEE Internet Computing*, *25*(5), 43–51. https://doi.org/10.1109/MIC.2021.3063257

Sayakhot, P., & Carolan-Olah, M. (2016). Internet use by pregnant women seeking pregnancy-related information: A systematic review. *BMC Pregnancy and Childbirth*, *16*, 65. https://doi.org/10.1186/s12884-016-0856-5

Schlicht, I. B., Fernandez, E., Chulvi, B., & Rosso, P. (2023). Automatic detection of health misinformation: A systematic review. *Journal of Ambient Intelligence and Humanized Computing*, 1–13. https://doi.org/10.1007/s12652-023-04619-4

Seah, S., & Weimann, G. (2020). What Influences the Willingness of Chinese WeChat Users to Forward Food-Safety Rumors? *International Journal of Communication*, *14*(0), Article 0.

Tang, X., Li, S., Gu, N., & Tan, M. (2019). Exploring repost features of police-generated microblogs through topic and sentiment analysis. *The Electronic Library*, *37*(4), 607–623. https://doi.org/10.1108/EL-02-2019-0044

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online | Science. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Xu, Q., Shen, Z., Shah, N., Cuomo, R., Cai, M., Brown, M., Li, J., & Mackey, T. (2020). Characterizing Weibo Social Media Posts From Wuhan, China During the Early Stages of the COVID-19 Pandemic: Qualitative Content Analysis. *JMIR Public Health and Surveillance*, *6*(4), e24125. https://doi.org/10.2196/24125

Zhang, S., Ma, F., Liu, Y., & Pian, W. (2021). Identifying features of health misinformation on social media sites: An exploratory analysis. *Library Hi Tech*, *40*(5), 1384–1401. https://doi.org/10.1108/LHT-09-2020-0242

Zhao, Y., Da, J., & Yan, J. (2021). Detecting health misinformation in online health communities: Incorporating behavioral features into machine learning based approaches. *Information Processing & Management*, *58*(1), 102390. https://doi.org/10.1016/j.ipm.2020.102390

Zhao, Y., Zhu, S., Wan, Q., Li, T., Zou, C., Wang, H., & Deng, S. (2022). Understanding How and by Whom COVID-19 Misinformation is Spread on Social Media: Coding and Network Analyses. *Journal of Medical Internet Research*, *24*(6), e37623. https://doi.org/10.2196/37623

Zhong, B. (2023). Going beyond fact-checking to fight health misinformation: A multi-level analysis of the Twitter response to health news stories. *International Journal of Information Management*, *70*, 102626. https://doi.org/10.1016/j.ijinfomgt.2023.102626

Zhou, C., Li, K., & Lu, Y. (2021). Linguistic characteristics and the dissemination of misinformation in social media: The moderating effect of information richness. *Information Processing & Management*, *58*(6), 102679. https://doi.org/10.1016/j.ipm.2021.102679

Zhu, Q. (2016). The application of social media in outreach of academic libraries' resources and services: A case study on WeChat. *Library Hi Tech*, *34*(4), 615–624. https://doi.org/10.1108/LHT-05-2016-0055