

Towards User Empowerment: Bridging the Gap in Health Misinformation Protection on Social Networks

HUIYUN TANG, University of Luxembourg, Luxembourg

VINCENT KOENIG, University of Luxembourg, Luxembourg

ANASTASIA SERGEEVA, University of Luxembourg, Luxembourg

Health misinformation in social networks requires immediate attention due to its severe consequences, as exemplified by the COVID-19 pandemic response on social media. However, the existing solutions designed to combat misinformation generally overlook the unique characteristics of health misinformation domain. Through a review of relevant literature and a critical analysis of current anti-misinformation solutions, we have identified significant user-side issues that undermine the effectiveness of existing approaches in addressing health misinformation. To tackle these issues, we put forth several strategies to empower users in combating health misinformation. Our research contributes to understanding the challenges associated with health misinformation correction on social networks.

Additional Key Words and Phrases: Misinformation, Health; Credibility Indicators, Design Frictions

Huiyun Tang, Vincent Koenig, and Anastasia Sergeeva. 2023. Towards User Empowerment: Bridging the Gap in Health Misinformation Protection on Social Networks. Position Paper at the Workshop "Designing for and Reflecting upon Resilience in Health and Wellbeing" at the ACM Designing Interactive Systems Conference (DIS'23), July 11, 2023, Pittsburgh, Pennsylvania, USA, 8 pages. <https://sites.psu.edu/dis2023workshophealthresilience/>

1 INTRODUCTION

In recent years, there has been a growing trend of people using social media to address their well-being and health concerns [36]. While social media health groups can provide broader community support to its users[47], they can also become breeding grounds for misinformation. The COVID-19 pandemic is a notable example of this phenomenon, with numerous Facebook groups and Twitter threads promoting extreme views and misinformation about the virus and its countermeasures. Studies have shown a significant impact of vaccine deniers and conspiracy Facebook groups on spreading misinformation about vaccines and hygiene practices among the general population [5, 20]. Several scholars have referred to the COVID-19 pandemic as an "infodemic" [31, 34], characterized by the rapid spread of information, including a substantial amount of false and misleading content. Social media platforms play a significant role in propagating such misinformation, leading people to act against public health policies and scientific guidance, ultimately causing harm to its users. Academics, government officials, industry stakeholders, and social media platform owners have suggested various ways of addressing the challenges posed by this infodemic [23, 28, 39]. Within the field of Human-Computer Interaction, current approaches to combat misinformation include design friction methods that aim to delay users from immediately sharing information, giving them time to evaluate the news before sharing [8, 19, 30]. Another line of research focuses on developing credibility indicators that signify potential credibility problems with content [13, 26, 44]. However, these solutions have shown limited effectiveness in combating misinformation [14, 42]. Furthermore, most current solutions do not consider health misinformation's specific characteristics, including the content's nature and users' motivations for sharing it. In this paper, we discuss how the characteristics of health misinformation limit the efficiency of existing anti-misinformation

Authors' addresses: Huiyun Tang, University of Luxembourg, Luxembourg, huiyun.tang@uni.lu; Vincent Koenig, University of Luxembourg, Luxembourg, vincent.koenig@uni.lu; Anastasia Sergeeva, University of Luxembourg, Luxembourg, anastasia.sergeeva@uni.lu.

solutions and how we can address these limitations to enhance the efficiency of credibility-building and misinformation prevention. Our proposals aim to make fact-checking and misinformation prevention elements in social network interfaces user-friendly, respectful of users' choices and preferences, and personalized to their needs. The future solutions aim to address users' doubts about the credibility of information sources, enable them to customize the information provided by different sources, and provide a tool for quick fact-checking before sharing information. To illustrate our ideas, we present a mock-up interface embedded in a fictional mobile Facebook plugin incorporating interactive features for fact-checking and misinformation prevention. We discuss the mock-ups and provide directions for future development and evaluation.

2 RELATED WORK

2.1 Characteristics of Health Misinformation in Social Networks

Health misinformation refers to health-related claims of fact that are currently considered to be false due to a lack of scientific evidence or contradicting the consensus of the scientific/expert community regarding a given topic [11, 37]. While most studies agree that health misinformation creates great risks to society [15, 17, 25], a limited number of studies discuss the specific characteristics of this type of misinformation in comparison with the characteristics of other types of misinformation (political or marketing) in relation to the interventions that can be applied to protect users from this specific kind of threat [42]. Recent studies indicate that when it comes to sharing information on social networks, the motivation behind sharing political misinformation is often driven by partisanship or enjoyment [27, 38]. However, for health-related topics, including both accurate information and misinformation, altruistic motivation emerges as one of the primary factors influencing sharing behavior [4, 46]. In relation to this, sharing health-related information in social networks often takes the form of personal stories, motivated by the search for people with shared experiences and seeking social support [35, 40] or, to some extent, for self-presentation purposes [29]. That means, in many cases, health-related information in social networks incorporates opinion-based features and personal experiences, which makes it difficult to be marked as "true" or "false," even if the mentioned treatments contradict expert consensus. If the system marks such content as false, it can be perceived as a personal attack by the individuals sharing their experiences. Another issue is that misleading health information sometimes includes genuine medical terms and concepts, which require professional expertise to fully understand and make appropriate content suggestions. [9]. In this context, the knowledge gap stops users from understanding the false claims within the text. Alternatively, even genuine medical information, when presented via social networks to non-professionals, can generate misinformation claims as non-professional users cannot fully understand the concepts and relations behind the information [24, 48].

2.2 Design Interventions and Solutions for Informed Decision-Making and Misinformation Reduction in Social Media Platforms

Previous research in the HCI domain has introduced various interventions to assist users in making informed decisions about the media they consume. These interventions enhance user awareness and credibility assessment through credibility rating scales, provenance indicators, warnings, and credibility labels. For instance, Enrico et al. incorporated a credibility rating scale and displayed the source of credibility in the user interface [7]. Amin et al. proposed a similar intervention using labels indicating credible fact-checking sources and pop-up warning notifications to capture users' attention [2]. Several works have proposed using provenance warnings to alert users when the source of information is unverified [12, 33]. Binary credibility labels, such as "thumbs up/down" icons, have also been utilized to indicate the integrity of news articles [6].

In addition to efforts focused on assessing information credibility, researchers have explored design interventions to raise awareness and promote critical news consumption behavior, aiming to reduce the spread of misinformation on social media [32]. These interventions can take the form of nudges and frictions. Nudges are designed to guide individuals in a specific direction without restricting their freedom of choice [16]. For example, Jahanbakhsh et al. developed behavioral nudges that provide users with accurate assessments and explanations regarding the accuracy of news stories, thereby reducing the dissemination of false content [19]. Similarly, Andi and Akesson investigated using "social norm-based nudges" to encourage individuals to share more responsibly. These nudges remind users that false news is prevalent online and that responsible individuals think twice before sharing information with friends and followers [3]. Chen et al. focused on nudging and boosting users to resist misinformation by exposing them to the presence of filter bubbles and presenting news beyond those bubbles [10]. Capraro and Celadin examined three types of nudges in the Facebook format and discovered that endorsing accuracy reduced the sharing of misinformation while increasing the sharing of genuine news [8].

An emerging group of solutions encourages users to engage with the content actively, offering them accessible tools to verify the authenticity of the information's source. The approach involves providing users with web browser extensions that offer quick access to information and context about specific publishers or content, enabling users to verify the credibility of information [1, 45] or providing clues about information credibility and offering an instrument to mark questionable content [21]. Interactive AI-based solutions have emerged as potential tools for combating misinformation, also offering users a more tangible way to interact with information. The approach was employed by Jahanbakhsh et al., who used a human-in-the-loop to present AI suggestions regarding possible misinformation in tweets and receive human feedback on the proposals. The system also allowed users to request explanations from the AI system regarding its decisions [18].

3 REFINING MISINFORMATION CORRECTION SOLUTIONS FOR HEALTH INFORMATION IN SOCIAL NETWORKS

While studies indicate that correcting misinformation in the health domain tends to be more successful than in the political sphere [43] and no significant evidence suggests that attempts correct health-related misinformation inadvertently increase people's acceptance of the falsehood [42], the overall impact of correction efforts remains in the weak-to-moderate range [42], which calls for a discussion on the potential limitations of current correction measures and the identification of additional factors that can enhance intervention effectiveness.

One criticism of current solutions is their **oversimplified approach to assessing message credibility**. The categorization of news as either "true" or "false" overlooks the intricate nuances and uncertainties. For instance, health-related information shared on social networks often takes the form of personal storytelling. Treating the entire narrative as "non-credible information" can be judgmental and impolite. The design community should explore better alternatives for addressing opinionated claims, such as developing tools to highlight specific pieces of information containing false claims while keeping the rest of the message intact. This approach would allow for a more nuanced understanding of the content.

The second critical limitation of current approaches to identifying non-factual content on social media, particularly in the context of health-related information, is the **lack of transparency**. In the health domain, users are likely to accept misinformation corrections when provided by credible organizations such as international health agencies[41]; studies show the high trust in medical scientists as a source of credible information [37]. Unlike political topics where source bias is more commonly perceived, users in health-related discussions are generally more open to reconsidering

their beliefs. Providing transparent access to external sources, especially trusted health sites, can enhance users' reliance on credibility indicators.

Lack of user agency can also be considered an important limitation in health-related domains. Most existing solutions in social networks passively provide users with final judgments about the integrity of information. However, research has shown that user agency is essential in combating online misinformation [22]. Previous studies have indicated a strong motivation for research and knowledge-seeking regarding sharing and seeking health-related information online [40]. To address this, we propose an interactive approach that actively engages users in asking questions about different aspects of health-related information. This approach can enhance their critical thinking and decision-making by empowering users to form their own opinions about the truthfulness of the information.

Finally, **health(mis)information often involves complex medical domain knowledge, making it challenging for users to comprehend**. To alleviate this issue, we suggest implementing a tool that can automatically summarize and simplify text while highlighting key concepts and explaining the main message using non-specialized language. This tool would not require specific medical knowledge, enabling users to understand the information better and make informed decisions about sharing it.

Based on the highlighted challenges and potential solutions, we have developed a series of mock-ups to illustrate these ideas. We aim to utilize these mock-ups in forthcoming discussions with users to explore the suitability, efficiency, and ethical considerations of current and future solutions. The mock-ups, along with their descriptions, are presented in the Appendix.

4 CONCLUSIONS

The prevalence of health misinformation in social networks requires immediate attention. Practitioners must recognize and comprehend the unique characteristics of this type of misinformation to develop specialized solutions. This paper serves as an initial step in bridging the gap between the problem of health misinformation and the broader field of misinformation-combating solutions, which are currently used in social networks. We hope that our work catalyzes future research endeavors that focus on designing more effective interventions in the health-related misinformation domain.

ACKNOWLEDGMENTS

This research is supported by the Luxembourg National Research Fund INTER/FNRS/21/16554939/REMEDIS

REFERENCES

- [1] 2023. Introducing: Newsguard. <https://www.newsguardtech.com/how-it-works/>
- [2] Zaid Amin, Nazlena Mohamad Ali, and Alan F Smeaton. 2021. Visual Selective Attention System to Intervene User Attention in Sharing COVID-19 Misinformation. *arXiv preprint arXiv:2110.13489* (2021).
- [3] Simge Andi and Jesper Akesson. 2020. Nudging away false news: Evidence from a social norms experiment. *Digital Journalism* 9, 1 (2020), 106–125.
- [4] Oberiri Destiny Apuke and Bahiyah Omar. 2021. User motivation in fake news sharing during the COVID-19 pandemic: an application of the uses and gratification theory. *Online Information Review* 45, 1 (2021), 220–239.
- [5] Philip Ball and Amy Maxmen. 2020. The epic battle against coronavirus misinformation and conspiracy theories. *Nature* 581, 7809 (2020), 371–375.
- [6] Ranojoy Barua, Rajdeep Maity, Dipankar Minj, Tarang Barua, and Ashish Kumar Layek. 2019. F-NAD: an application for fake news article detection using machine learning techniques. In *2019 IEEE Bombay section signature conference (IBSSC)*. IEEE, 1–6.
- [7] Enrico Bunde, Niklas K uhl, and Christian Meske. 2021. Fake or Credible? Towards Designing Services to Support Users' Credibility Assessment of News Content. *arXiv preprint arXiv:2109.13336* (2021).

- [8] Valerio Capraro and Tatiana Celadin. 2022. “I think this news is accurate”: Endorsing accuracy decreases the sharing of fake news and increases the sharing of real news. *Personality and Social Psychology Bulletin* (2022), 01461672221117691.
- [9] Canyu Chen, Haoran Wang, Matthew Shapiro, Yunyu Xiao, Fei Wang, and Kai Shu. 2022. Combating Health Misinformation in Social Media: Characterization, Detection, Intervention, and Open Issues. *arXiv preprint arXiv:2211.05289* (2022).
- [10] Guangyu Chen, Paolo Ciuccarelli, and Sara Colombo. 2022. VisualBubble: Exploring How Reflection-Oriented User Experiences Affect Users’ Awareness of Their Exposure to Misinformation on Social Media. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–7.
- [11] Wen-Ying Sylvia Chou, April Oh, and William MP Klein. 2018. Addressing health-related misinformation on social media. *Jama* 320, 23 (2018), 2417–2418.
- [12] Frans Folkvord, Freek Snelling, Doeschka Anschutz, Tilo Hartmann, Alexandra Theben, Laura Gunderson, Ivar Vermeulen, and Francisco Lupiáñez-Villanueva. 2022. Effect of source type and protective message on the critical evaluation of news messages on Facebook: Randomized controlled trial in the Netherlands. *Journal of Medical Internet Research* 24, 3 (2022), e27945.
- [13] Dilrukshi Gamage, James Stomber, Farnaz Jahanbakhsh, Bill Skeet, and Gautam Kishore Shahi. 2022. Designing Credibility Tools To Combat Mis/Disinformation: A Human-Centered Approach. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–4.
- [14] Mingkun Gao, Ziang Xiao, Karrie Karahalios, and Wai-Tat Fu. 2018. To label or not to label: The effect of stance and credibility labels on readers’ selection and perception of news articles. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–16.
- [15] Amira Ghenai and Yelena Mejova. 2018. Fake cures: user-centric modeling of health misinformation in social media. *Proceedings of the ACM on human-computer interaction* 2, CSCW (2018), 1–20.
- [16] Ralph Hertwig and Till Grüne-Yanoff. 2017. Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science* 12, 6 (2017), 973–986.
- [17] Shandell Houlden, Jaigris Hodson, George Veletsianos, Darren Reid, and Chris Thompson-Wagner. 2021. The health belief model: How public health can address the misinformation crisis beyond COVID-19. *Public Health in Practice* 2 (2021), 100151.
- [18] Farnaz Jahanbakhsh, Yannis Katsis, Dakuo Wang, Lucian Popa, and Michael Muller. 2023. Exploring the Use of Personalized AI for Identifying Misinformation on Social Media. (2023).
- [19] Farnaz Jahanbakhsh, Amy X Zhang, Adam J Berinsky, Gordon Pennycook, David G Rand, and David R Karger. 2021. Exploring lightweight interventions at posting time to reduce the sharing of misinformation on social media. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–42.
- [20] Seth C Kalichman, Lisa A Eaton, Valerie A Earnshaw, and Natalie Brousseau. 2022. Faster than warp speed: early attention to COVID-19 by anti-vaccine groups on Facebook. *Journal of Public Health* 44, 1 (2022), e96–e105.
- [21] Yash Khivasara, Yash Khare, and Tejas Bhadane. 2020. Fake news detection system using web-extension. In *2020 IEEE Pune Section International Conference (PuneCon)*. IEEE, 119–123.
- [22] Eleni A Kyza and Christiana Varda. 2019. NAVIGATING THE POST-TRUTH ERA: TRUST, MISINFORMATION, AND CREDIBILITY ASSESSMENT ON ONLINE SOCIAL MEDIA. *AoIR Selected Papers of Internet Research* (2019).
- [23] Junmin Lee, Keungoui Kim, Gangmin Park, and Namjun Cha. 2021. The role of online news and social media in preventive action in times of infodemic from a social capital perspective: The case of the COVID-19 pandemic in South Korea. *Telematics and Informatics* 64 (2021), 101691.
- [24] Jisan Lee, Jongkwan Koh, and Jong-Yeup Kim. 2021. Popularization of medical information. *Healthcare Informatics Research* 27, 2 (2021), 110–115.
- [25] Yang-Jun Li, Christy MK Cheung, Xiao-Liang Shen, and Matthew KO Lee. 2019. Health misinformation on social media: a literature review. (2019).
- [26] Zhuoran Lu, Patrick Li, Weilong Wang, and Ming Yin. 2022. The Effects of AI-based Credibility Indicators on the Detection and Spread of Misinformation under Social Influence. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–27.
- [27] Miriam J Metzger, Andrew J Flanagin, Paul Mena, Shan Jiang, and Christo Wilson. 2021. From dark to light: The many shades of sharing misinformation online. *Media and Communication* 9, 1 (2021), 134–143.
- [28] Azzam Mourad, Ali Srour, Haidar Harmanani, Cathia Jenainati, and Mohamad Arafeh. 2020. Critical impact of social networks infodemic on defeating coronavirus COVID-19 pandemic: Twitter-based study and research directions. *IEEE Transactions on Network and Service Management* 17, 4 (2020), 2145–2155.
- [29] Mark W Newman, Debra Lauterbach, Sean A Munson, Paul Resnick, and Margaret E Morris. 2011. It’s not that I don’t have problems, I’m just not putting them on Facebook: challenges and opportunities in using online social networks for health. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. 341–350.

- [30] Folco Panizza, Piero Ronzani, Carlo Martini, Simone Mattavelli, Tiffany Morisseau, and Matteo Motterlini. 2022. Lateral reading and monetary incentives to spot disinformation about science. *Scientific Reports* 12, 1 (2022), 5678.
- [31] Parth Patwa, Shivam Sharma, Srinivas Pykl, Vineeth Guptha, Gitanjali Kumari, Md Shad Akhtar, Asif Ekbal, Amitava Das, and Tanmoy Chakraborty. 2021. Fighting an infodemic: Covid-19 fake news dataset. In *Combating Online Hostile Posts in Regional Languages during Emergency Situation: First International Workshop, CONSTRAINT 2021, Collocated with AAAI 2021, Virtual Event, February 8, 2021, Revised Selected Papers 1*. Springer, 21–29.
- [32] Emily Saltz, Tommy Shane, Victoria Kwan, Claire Leibowicz, and Claire Wardle. 2020. It matters how platforms label manipulated media. Here are 12 principles designers should follow.-The Partnership on AI.
- [33] Imani N Sherman, Jack W Stokes, and Elissa M Redmiles. 2021. Designing Media Provenance Indicators to Combat Fake Media. In *Proceedings of the 24th International Symposium on Research in Attacks, Intrusions and Defenses*. 324–339.
- [34] Daniel H Solomon, Richard Bucala, Mariana J Kaplan, and Peter A Nigrovic. 2020. The “Infodemic” of COVID-19. , 1806–1808 pages.
- [35] H Song, K Komori, J Kim, K Tenzek, J Hawkins, W Lin, and JY Jung. 2016. Trusting social media as a source of health information: Comparison among US, Korea, and Hong Kong. *Journal of Medical Internet Research* 18, 3 (2016).
- [36] Victor Suarez-Lledo and Javier Alvarez-Galvez. 2021. Prevalence of health misinformation on social media: systematic review. *Journal of medical Internet research* 23, 1 (2021), e17187.
- [37] Briony Swire-Thompson, David Lazer, et al. 2020. Public health and online misinformation: challenges and recommendations. *Annu Rev Public Health* 41, 1 (2020), 433–451.
- [38] Wee-Kheng Tan and Chun Yu Hsu. 2023. The application of emotions, sharing motivations, and psychological distance in examining the intention to share COVID-19-related fake news. *Online Information Review* 47, 1 (2023), 59–80.
- [39] Viroj Tangcharoensathien, Neville Calleja, Tim Nguyen, Tina Purnat, Marcelo D’Agostino, Sebastian Garcia-Saiso, Mark Landry, Arash Rashidian, Clayton Hamilton, Abdelhalim AbdAllah, et al. 2020. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. *Journal of medical Internet research* 22, 6 (2020), e19659.
- [40] Sadegh Torabi and Konstantin Beznosov. 2016. Sharing health information on facebook: practices, preferences, and risk perceptions of North American users. In *Symposium on Usable Privacy and Security (SOUPS)*.
- [41] Emily K Vraga and Leticia Bode. 2017. Using expert sources to correct health misinformation in social media. *Science communication* 39, 5 (2017), 621–645.
- [42] Nathan Walter, John J Brooks, Camille J Saucier, and Sapna Suresh. 2021. Evaluating the impact of attempts to correct health misinformation on social media: A meta-analysis. *Health Communication* 36, 13 (2021), 1776–1784.
- [43] Nathan Walter and Sheila T Murphy. 2018. How to unring the bell: A meta-analytic approach to correction of misinformation. *Communication monographs* 85, 3 (2018), 423–441.
- [44] Waheeb Yaqub, Otari Kakhidze, Morgan L Brockman, Nasir Memon, and Sameer Patil. 2020. Effects of credibility indicators on social media news sharing intent. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–14.
- [45] Amy X Zhang, Aditya Ranganathan, Sarah Emlen Metz, Scott Appling, Connie Moon Sehat, Norman Gilmore, Nick B Adams, Emmanuel Vincent, Jennifer Lee, Martin Robbins, et al. 2018. A structured response to misinformation: Defining and annotating credibility indicators in news articles. In *Companion Proceedings of the The Web Conference 2018*. 603–612.
- [46] Xing Zhang, Shan Liu, Zhaohua Deng, and Xing Chen. 2017. Knowledge sharing motivations in online health communities: A comparative study of health professionals and normal users. *Computers in Human Behavior* 75 (2017), 797–810.
- [47] Yan Zhang, Dan He, and Yoonmo Sang. 2013. Facebook as a platform for health information and communication: a case study of a diabetes group. *Journal of medical systems* 37 (2013), 1–12.
- [48] Yuehua Zhao, Jingwei Da, and Jiaqi Yan. 2021. Detecting health misinformation in online health communities: Incorporating behavioral features into machine learning based approaches. *Information Processing & Management* 58, 1 (2021), 102390.

A APPENDIX: MOCK-UPS OF PROPOSED SOLUTIONS

An app presented on Figure 1 aims to resolve the issues linked to oversimplified news credibility and the absence of transparency in generating credibility indicators. The goal is to provide users with clear information about the source of opinions regarding the reliability of news articles. Frame 1 shows the initial news feed. The second frame explains our approach to viewing information as a mixture of facts that require verification rather than simply true or false. We also emphasize that it is the user’s responsibility to assess the credibility of the information. In Frames 4 and 5,

we highlight questionable information and include links to credible sources that can verify the information.

Figure 2 presents a way to empower users by giving them agency and providing a toolbox of solutions to actively transform information, making it easier to access and evaluate. The first frame represents the initial text, which can be challenging to understand and verify. To interact with the text, we allow users to select a specific part of the text and ask the system a question related to that particular section (Frame 2). Another option is to utilize a module that simplifies the text (Frames 3-4), offering a quick overview of the information by summarizing it (Frames 5-6).

Fig. 1

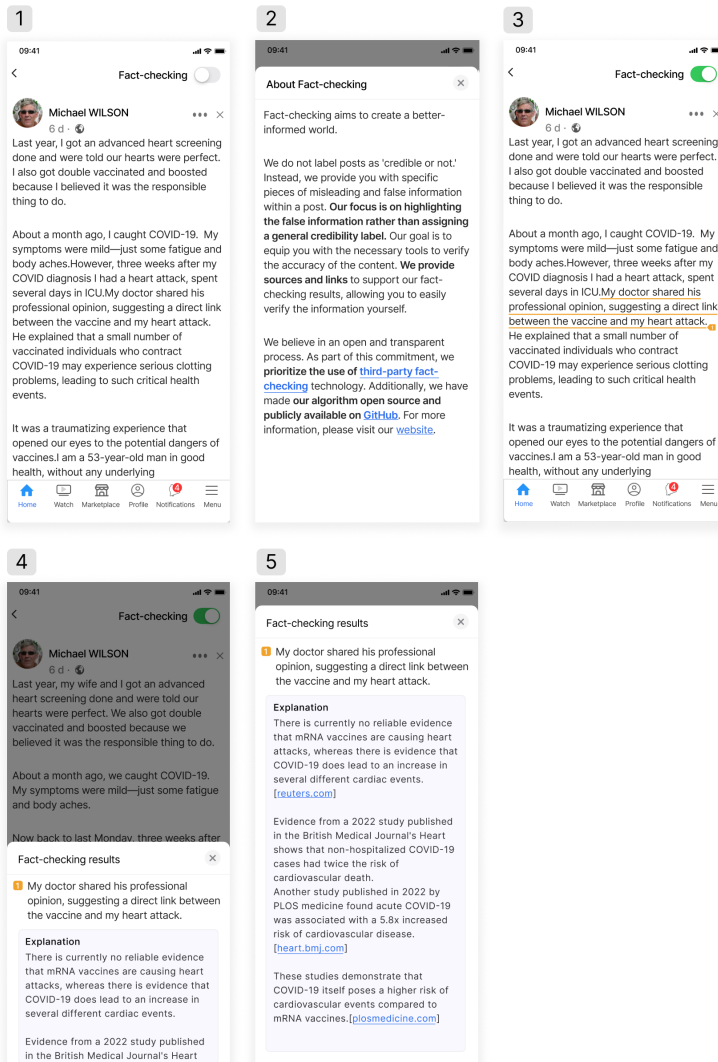


Fig. 2

