Mitigating Misinformation in User-Generated Discourse

Khawar Murad Ahmed
Kahlert School of Computing
University of Utah
Salt Lake City, Utah, United States
khawar.ahmed@utah.edu

Abstract

Interactive content in online spaces is often meant to inform a broad audience regarding issues of societal interest. Recently, combating misinformation has emerged as another task requiring interventions to reach broad audiences. We are engaged in two efforts to investigate how lay users leverage social media to inform and engage a broader audience and the characteristics of the resulting user-generated discourse. We found that tailoring content for broader audiences can be laborious and the impact of the efforts is neither guaranteed nor immediately apparent. Our future work will focus on approaches to help lay users mitigate these problems.

CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**; **Empirical studies in collaborative and social computing**; • **Information systems** → Social networks.

Keywords

user-generated content, online discourse, online communities, social media, misinformation, fact-checking

ACM Reference Format:
Khawar Murad Ahmed. 2025. Mitigating Misinformation in User-Generated Discourse. In *The 2025 ACM International Conference on Supporting Group Work (GROUP Companion '25), January 12–15, 2025, Hilton Head, SC, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3688828.3699660

1 Introduction

Social media users interact on the platforms and discuss a range of issues. These interactions often involve specific issues of societal interest, such as ongoing conflicts, natural disasters, education, public health, etc. The interactive content produces a user-generated discourse that can inform a broad audience beyond the parties involved in direct communication with each other. For example, the widespread use of #MeToo led many women to share their stories regarding sexual harassment and assault, demonstrating the scale of the problem and its systemic underpinning [5].

In recent times, combating online misinformation has emerged as another task that requires interventions that reach a broad audience. One such intervention is to provide fact-checks for content containing misinformation. However, manual fact-checking by professionals is slow and does not scale [11] while automated means of fact-checking typically lack a nuanced understanding of the context needed to make accurate fact-checking judgments [6]. User-generated discourse could provide a possible means to overcome these shortcomings. For instance, the CoFact service in Taiwan leveraged volunteers to help debunk rumors circulating in the popular messaging app, LINE [10].

In order to enrich our understanding of how lay users leverage social media to inform and engage a broader audience, we focus on the following research questions:

- What are the practices of lay users who generate content aimed at informing and engaging a broader audience through their communication? (*User focus*)
- What are the characteristics of user-generated discourse aimed at informing and engaging broader audiences? (*Content focus*)

By addressing the research questions above, we will gather insight that supports lay users in reaching and accurately informing their intended audiences and enhance the overall societal utility of user-generated content. Networking with peers and experienced researchers in the GROUP 2025 Doctoral Consortium can develop and shape the research efforts to help achieve this goal.

2 Related Work

Our work is inspired by two strands of literature. First, our work is related to the literature that examines user-generated discourse on social media through the lens of online communities. Second, our work builds on the literature that points to the ability of lay users to detect and correct misinformation. We summarize the salient literature each of the two strands below.

2.1 User-generated Discourse in Online Communities

Social media platforms serve as a means to the formation of various online communities. The features of the platform can influence the formation of these communities and the discourse among users within the communities. For example, discourse on Reddit takes place in specific topic-based communities called subreddits [15]. On the other hand, X (formerly Twitter) communities typically coalesce around a hashtag to discuss and amplify certain issues. Such communities are akin to what Bruns and Burgess [1] has referred to as"ad-hoc publics." For example, users posted tweets using the hashtag #fridaysforfuture to demand action from political leaders to address the issues raised by climate change [8]. Increasingly, online communities are affected by content containing misinformation that may either be posted by users within the community [12] or

GROUP Companion '25, January 12–15, 2025, Hilton Head, SC, USA

ACM ISBN 979-8-4007-1187-9/25/01
https://doi.org/10.1145/3688828.3699660

injected deliberately by disinformation campaigns that typically leverage bots [16].

2.2 Crowdsourced Fact-checking

Given the severity of online misinformation, researchers have studied several interventions to combat misinformation. One such intervention is leveraging the crowd in various ways. Several studies demonstrate that laypeople can be as effective as seasoned journalists at detecting misinformation [13]. Additionally, features, such as Community Notes [1] on X, that involve fact-checking by ordinary users have been shown to be effective at countering false claims [4, 7].

The literature also provides some insight regarding the motivations of lay people who engage in fact-checking online. For instance, Drolsbach and Pröllochs [3] suggest that laypeople opt to fact-check misinformation that garners higher engagement from the community members. Other studies indicate that lay users often select a claim to fact-check based on two factors: relationship with the poster [14] and interest in the issue [2].

3 Work in Progress

We are addressing the research questions listed in Section 1 with two research efforts, each exploring different aspects of user-generated discourse. One of these characterizes the user-generated discourse that was targeted specifically at Western audiences during the Ukraine-Russia conflict in 2014 (see Section 3.1). The other is aimed at understanding the experiences of laypeople who correct misinformation they encounter within online communities they inhabit (see Section 3.2).

3.1 External-Facing User-Generated Discourse during the Ukraine-Russia conflict in 2014

During the Ukraine-Russia conflict in 2014, users on each side of the conflict turned to Twitter to make their respective cases to people in Western countries. We are analyzing such externally-facing user-generated discourse during this conflict by collecting Twitter data based on relevant keywords.

We classified individual tweets within the dataset into one of three categories: those supporting each side of the conflict and those that were neutral. In addition, we examined the URLs (Uniform Resource Locators) included within the tweets in each category. Our initial analysis indicates that users specifically crafted the tweets for an audience external to the region of the conflict. For instance, the URLs included within the tweets in order to bolster the credibility of the message tend to link to Western media (e.g., the New York Times). The above observations suggest the posters were keenly aware of the need to craft their messages in ways that might be appealing to the target audience.

3.2 Fact-checking by Lay Users

Given the recent explosion of online misinformation within online communities, we conducted an interview study consisting of lay users who reported having corrected online misinformation. Our initial findings reveal that some lay users are highly motivated to offer fact-checks to counter the misinformation they encounter in online communities. These fact-checks are meant to serve a broader audience beyond the originator of the misinformation. Specifically, the lay users who engage in fact-checking do so to mitigate the harmful impacts of misinformation and to raise greater awareness of factual information on the issue in question.

At the same time, our data suggests that lay users lack a shared standard for systematically analyzing misinformation and communicating fact-checks. Moreover, uncertainty regarding the impact of their fact-checking on the broader audience can lead them to limit their efforts or stop fact-checking altogether.

4 Discussion and Future Work

Our current work suggests that lay users tailor the content they share with their target audiences using various platform features. This is exemplified by the greater use of Western news sources in externally-facing discourse about the Ukraine-Russia conflict (see Section 3.1). We additionally found that lay users care about the broader impact of their activities. For instance, the lay users who engage in fact-checking care about countering the harmful societal impacts of misinformation (see Section 3.2).

Given that media organizations usually have broader reach and greater credibility, our findings suggest that there is an opportunity for collaboration between laypeople and domain experts. The literature on citizen science provides numerous examples of successful collaborations between laypeople and experts, such as the Foldit crowdsourcing platform for biochemists to engage ordinary citizens in their protein research [9]. Such types of collaboration could potentially be applied to help professional fact-checkers in combating online misinformation while providing lay users the opportunity to reach their target audiences more effectively.

However, a direct collaboration between laypeople and media professionals is likely to run into at least three challenges. First, the distribution of effort might be affected by the reputation of lay users not being high enough to garner the same level of algorithmic boost and visibility as the content posted by renowned individuals. Second, the collaboration may require that the crowd receive basic training on informative online content production. Third, the collaborative arrangements may face access hurdles because a large volume of user-generated discourse is not publicly available because it takes place in non-public groups or direct messages between users. It is likely that no single solution could effectively addresses the three challenges simultaneously.

As the first step toward overcoming the above challenges, we need to understand how domain experts leverage the efforts of lay users to reach broad audiences. By understanding the interaction between domain experts and lay users, we hope to gather insight that enables better collaboration between the parties to facilitate more productive and useful discourse within online spaces.

5 Conclusion

We explored the motivations and practices of lay users for communicating with broader audiences in online communities. In addition, we examined the discursive characteristics of contributions and the potential impact on their intended audience. Our initial findings suggest that generating content for broader audiences can be

[1] https://communitynotes.x.com/guide/en/about/introduction

laborious and the impact of the efforts is neither guaranteed nor immediately apparent. Our future work will focus on approaches to help lay users mitigate these problems. Addressing these issues has the potential to promote more informative and impactful discourse in online communities.

Acknowledgments

I thank our collaborators Christopher de Freitas and Sarah Choe for their contributions to the research efforts described in this paper. The work described in this paper is carried out under the valuable guidance of my advisors Marina Kogan and Sameer Patil. Parts of the research work are supported by the US National Science Foundation (NSF) grant #2154123. The contents of the paper do not necessarily reflect the views of the sponsors.

References

[1] Axel Bruns and Jean Burgess. 2011. The use of Twitter hashtags in the formation of ad hoc publics. In *Proceedings of the 6th European Consortium for Political Research (ECPR) General Conference 2011*, A Bruns and P De Wilde (Eds.). The European Consortium for Political Research (ECPR), United Kingdom, 1–9. https://eprints.qut.edu.au/46515/

[2] Elizabeth L. Cohen, Anita Atwell Seate, Stephen M. Kromka, Andrew Sutherland, Matthew Thomas, Karissa Skerda, and Andrew Nicholson. 2020. To correct or not to correct? Social identity threats increase willingness to denounce fake news through presumed media influence and hostile media perceptions. *Communication Research Reports* 37, 5 (2020), 263–275. https://doi.org/10.1080/08824096.2020.1841622

[3] Chiara Patricia Drolsbach and Nicolas Pröllochs. 2023. Diffusion of Community Fact-Checked Misinformation on Twitter. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 267 (oct 2023), 22 pages. https://doi.org/10.1145/3610058

[4] Chiara Patricia Drolsbach, Kirill Solovev, and Nicolas Pröllochs. 2024. Community notes increase trust in fact-checking on social media. *PNAS Nexus* 3, 7 (05 2024), 217–230. https://doi.org/10.1093/pnasnexus/pgae217

[5] Ryan J. Gallagher, Elizabeth Stowell, Andrea G. Parker, and Brooke Foucault Welles. 2019. Reclaiming Stigmatized Narratives: The Networked Disclosure Landscape of #MeToo. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 96 (Nov. 2019), 30 pages. https://doi.org/10.1145/3359198

[6] D Graves. 2018. Understanding the promise and limits of automated fact-checking. *Reuters Institute for the Study of Journalism* (2018).

[7] Naeemul Hassan, Mohammad Yousuf, Md Mahfuzul Haque, Javier A. Suarez Rivas, and Md Khadimul Islam. 2019. Examining the Roles of Automation, Crowds and Professionals Towards Sustainable Fact-checking. In *Companion Proceedings of The 2019 World Wide Web Conference* (San Francisco, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 1001–1006. https://doi.org/10.1145/3308560.3316734

[8] Jörg Haßler, Anna-Katharina Wurst, Marc Jungblut, and Katharina Schlosser. 2023. Influence of the pandemic lockdown on Fridays for Future's hashtag activism. *New Media & Society* 25, 8 (2023), 1991–2013. https://doi.org/10.1177/14614448211026575

[9] Robert Kleffner, Jeff Flatten, Andrew Leaver-Fay, David Baker, Justin B Siegel, Firas Khatib, and Seth Cooper. 2017. Foldit Standalone: a video game-derived protein structure manipulation interface using Rosetta. *Bioinformatics* 33, 17 (05 2017), 2765–2767. https://doi.org/10.1093/bioinformatics/btx283

[10] Mei-chun Lee. 2024. Checking Facts by a Bot: Crowdsourced Facts and Intergenerational Care in Posttruth Taiwan. *Current Anthropology* 65, 4 (2024), 653–673.

[11] Nicholas Micallef, Vivienne Armacost, Nasir Memon, and Sameer Patil. 2022. True or False: Studying the Work Practices of Professional Fact-Checkers. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 127 (apr 2022), 44 pages. https://doi.org/10.1145/3512974

[12] Pardis Pourghomi, Milan Dordevic, and Fadi Safieddine. 2018. The Spreading of Misinformation online: 3D Simulation. In *2018 5th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*. 299–304. https://doi.org/10.1109/ICITACEE.2018.8576937

[13] Paul Resnick, Aljohara Alfayez, Jane Im, and Eric Gilbert. 2023. Searching for or reviewing evidence improves crowdworkers' misinformation judgments and reduces partisan bias. *Collective Intelligence* 2, 2 (2023), 1–15. https://doi.org/10.1177/26339137231173407

[14] Lauren Scott, Lynne Coventry, Marta E. Cecchinato, and Mark Warner. 2023. "I figured her feeling a little bit bad was worth it to not spread that kind of hate": Exploring how UK families discuss and challenge misinformation. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 660, 15 pages. https://doi.org/10.1145/3544548.3581202

[15] Ahmed Soliman, Jan Hafer, and Florian Lemmerich. 2019. A Characterization of Political Communities on Reddit. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media* (Hof, Germany) *(HT '19)*. Association for Computing Machinery, New York, NY, USA, 259–263. https://doi.org/10.1145/3342220.3343662

[16] Kate Starbird, Ahmer Arif, and Tom Wilson. 2019. Disinformation as Collaborative Work: Surfacing the Participatory Nature of Strategic Information Operations. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 127 (nov 2019), 26 pages. https://doi.org/10.1145/3359229