The Role of Artificial Intelligence in Combating Climate Change Misinformation Across Digital Media Platforms

Sheikh Fakhar e Alam [1] Aysal Elham [2] Madiha Qamar [3] Akash Ali [4]

[1] PhD Media & Communication Studies, Communication Specialist, Elementary & Secondary Education Department, Govt. of Khyber Pakhtunkhwa, Pakistan. Email: fakharhrp@gmail.com

[2] Independent Researcher, Pakistan. Email: aysalelham@gmail.com

[3] Creative Communication Consultant, Department of Communication, Riphah International University, Islamabad, Pakistan. Email: maditechnest@gmail.com

[4] MS Media Studies, Department of Communication, Riphah International University, Islamabad, Pakistan. Email: aa0408500@gmail.com

ABSTRACT: Climate-change misinformation on digital media platforms is harming the public understanding of environmental risk and slowing collective action to tackle climate change. Artificial intelligence (AI) systems are becoming a potential frontline of defence against such misinformation. This study explores the perceived impact of AI-assisted solutions such as automated fact-checking, natural language processing (NLP) content moderation, deepfake detection, algorithmic recommender re-ranking, and chatbot inoculation on decreasing exposure to and trust in climate misinformation on major social media platforms. A cross-sectional survey using a quantitative method was used. A structured questionnaire was administered on 300 active digital media users who were selected using stratified random sampling technique. Descriptive statistics, paired-sample t-tests and one-way ANOVA were used to analyse the data in SPSS v29. The mean effectiveness was highest ($M = 4.32$, $SD = 0.71$) for automated fact-checking, on a five-point Likert scale. Exposure to AI interventions significantly reduced weekly exposure to climate misinformation by an average of 44.6 % ($t = 12.87$, $p < .001$). Among the respondents, 68% trusted the AI-verified climate content to a moderate to high degree. Though there are challenges to broad-scale adoption, such as algorithmic transparency and cultural adaptability, AI-based interventions have the potential to greatly diminish exposure to climate misinformation and improve public trust in reliable information.

KEYWORDS: Artificial Intelligence, Climate Change, Misinformation, Digital Media, Fact-Checking, Machine Learning, Social Media Platforms

Pages: 270 – 279
Volume: 5
Issue: 5 (May 2026)

Corresponding Author
Sheikh Fakhar e Alam
✉ fakharhrp@gmail.com

Cite this Article: Alam, S. F. e., Elham, A., Qamar, M., & Ali, A. (2026). The Role of Artificial Intelligence in Combating Climate Change Misinformation Across Digital Media Platforms. *The Regional Tribune, 5*(5), 270-279. https://doi.org/10.55737/trt/v-v.300

Introduction

Climate change is one of the greatest challenges of the 21st century, but information overload continues to shape public opinion and action. Digital media platforms are the main information environment for billions of people, and at the same time they are amplifiers of unverified, misinformed, and intentionally misleading climate

information (Treen et al., 2020). The dissemination of these falsehoods to the public, influencing public opinion and thereby deferring meaningful climate action, continues at a wide scale (Lewandowsky, 2021). The volume, velocity and individualization of digital communication have rendered conventional methods of correction, like editorial fact-checking and expert opinion, incomplete.

In the present context, AI has come out as a significant player in the fight against misinformation on climate change. Machine-learning classifiers are capable of identifying a contrarian claim with a high level of accuracy through large volumes of textual material (Coan et al., 2021); natural language processing pipelines can be used to identify and demote harmful posts, and contextualize them when they are initially posted (Zhou & Zafarani, 2020). The AI toolkit is also complemented by additional innovations, including deepfake detectors, recommender re-ranking algorithms and conversational agents, which are designed to pre-emptively inoculate users against manipulation (Roozenbeek & van der Linden, 2018). Such technologies can assist in reducing misinformation as a community, and, in fact, improve the evidence-based climate discussion in the popular press.

However, the use of AI to combat climate misinformation is far from simple. However, users are wary of automated moderation, fearing censorship, opacity of the algorithms, and cultural bias (Pennycook & Rand, 2019). Misinformation is known to spread faster than truth on social media platforms such as Facebook, X (previously Twitter), and YouTube (Vosoughi et al., 2018), while co-ordinated botnets are still being used to spread misinformation to influence environmental debates (Shao et al., 2018). Therefore, it is important to gain a solid understanding of the performance of AI systems in combating misinformation and the user experience of such systems, to support evidence-informed policies and platform design (Ullah, 2026).

This research builds on this growing body of literature by conducting an empirical study with 300 active digital media platform users to explore their perceptions of the effectiveness of AI-driven responses to combat climate change misinformation. It also examines how much such interventions lead to a change in self-reported exposure to false content, and how they affect the trust placed in environmental information provided by AI. The study aims to deliver descriptive and inferential evidence that can be used practically, replicably, and in the context of platform governance, environmental communication, and public information policy in the digital age (Falkenberg et al., 2022).

Problem Statement

With AI systems being used across digital media platforms, climate change misinformation has spread to billions of people every day, misinforming the public and slowing policymaking. Existing studies have a limited amount of empirical evidence on user perceptions of AI-based countermeasures and the effectiveness of these tools to limit exposure to misinformation. This disconnection is a limitation on evidence-informed decisions related to platform governance, environmental communication, and the responsible integration of AI into public information ecosystems globally.

Literature Review

Climate-Change Misinformation on Digital Media

Digital platforms are the primary platforms for the production and propagation of climate misinformation. As Treen et al. (2020) have noted, networked platforms can amplify counter-narratives, with emotional framing and visuals, to a global audience in a short period. Indeed, false news spreads significantly faster and further than true news on X (Twitter), and the category of the environment is one of the most virally distorted categories, as found by Vosoughi

et al. (2018). Moreover, Farrell et al. (2019) also reported on "organized denial networks" working to normalize false claims in the public discourse.

Artificial Intelligence Approaches to Misinformation Detection

The recent developments in machine learning and NLP have revolutionized misinformation detection. Coan et al. (2021) built a classifier that could classify adverse climate claims in large text corpora with greater than 90 % accuracy. Zhou and Zafarani (2020) presented a detailed taxonomy of fake-news detection models, which highlights the increasing importance of transformer-based architectures. In addition to detection, Shao et al (2018) demonstrated automated systems' ability to detect low-credibility content disseminated via bot networks, and Pennycook and Rand (2019) illustrated how crowd-sourced signals can aid algorithmic judgments' reliability.

User Trust and Inoculation Strategies

Transparency and fairness are key factors to gain public trust in AI-verified content. Van der Linden et al. (2017) showed that brief exposures to weak misinformation boost resistance to subsequent misinformation attempts, which has been incorporated into AI chatbots (Roozenbeek & van der Linden, 2018). Similarly, Cook et al. (2017) demonstrated that factual corrections coupled with explanations of manipulation tactics removed misinformation more effectively than corrections alone. In a study by Diehl et al. (2021), the authors found that trust in social media is a significant predictor of trust in climate content, highlighting the need for trustworthy AI mediation.

Research Gap

While the literature confirms both the scale of misinformation and the capabilities of AI-based countermeasures, few studies have quantitatively assessed how ordinary users perceive the effectiveness of specific AI tools in the context of climate change (Ullah et al., 2024). Falkenberg et al. (2022) highlight rising polarisation on climate topics, yet the role of user-facing AI interventions in mitigating that polarisation remains under-examined. The present study addresses this gap.

Research Objectives and Questions

Research Objectives

1. To examine the extent to which AI-driven interventions reduce user exposure to climate-change misinformation across digital media platforms.
2. To evaluate the perceived effectiveness of specific AI tools (fact-checking, NLP moderation, deepfake detection, recommender re-ranking, and chatbot inoculation) in combating climate misinformation.
3. To determine the level of user trust in AI-verified climate information and its variation across demographic groups.
4. To identify practical implications for platform governance and environmental communication policy.

Research Questions

RQ1: To what extent do AI-driven interventions reduce user exposure to climate-change misinformation on digital media platforms?

RQ2: Which users perceive AI-based tools as most effective in detecting and countering climate misinformation?

RQ3: How does user trust in AI-verified climate information differ across age, education, and platform-use categories?

RQ4: What are the policy and platform-design implications of user-facing AI moderation for climate communication?

Research Methodology

Research Design

The study used a quantitative approach and a survey-type cross-sectional study with a positivist paradigm. This design was selected as it will enable the systematic collection of users' perceptions and behavioural outcomes in a large sample for statistical generalisation to similar populations of digital media users.

Population and Sampling

The respondents were active digital media users (18+ years) with active accounts on at least one of the major platforms (Facebook, X, YouTube, TikTok, Instagram, WhatsApp). Stratified random sampling was used to ensure age- and gender-balance and platform preference. Using Cochran's formula, the final sample size was computed at a 95% confidence level with a 5.6% margin of error for 300 respondents.

Research Instrument

This questionnaire was a structured online survey that comprised four questions: (a) demographic profile, (b) exposure to climate misinformation, (c) perceived effectiveness of AI-based interventions, and (d) trust in AI-verified content. Perceptual items were rated using a 5-point Likert scale (1 = strong disagree; 5 = strong agree). The pilot test provided internal consistency, with 25 respondents yielding Cronbach's 0.87.

Data Collection Procedure

The survey was sent electronically via Google Forms and posted in platform communities over 6 weeks. Informed consent was obtained, anonymity was preserved, and participation was voluntary. The overall response rate was 82%.

Data Analysis

The data were gathered and processed with the help of the software IBM SPSS Statistics v29. The responses were summarised using descriptive statistics (frequencies, means and standard deviations), and inferential statistical testing (paired-sample t-tests and one-way ANOVA) was conducted to identify the differences before and after exposure to the AI-based interventions, and between demographic groups. The statistically significant differences were established at $p<0.05$.

Ethical Considerations

Ethical considerations were adhered to and involved informed consent, data anonymity, and safe storage of data on encrypted servers.

Findings of the Study

Demographic Profile

Table 1 summarises demographic data of the 300 respondents. The sample was balanced in both sexes (52 % female, 48 percent male) and a broad age and educational spectrum that gives it a broad interpretive spectrum.

Table 1

Demographic Characteristics of Respondents (N = 300)

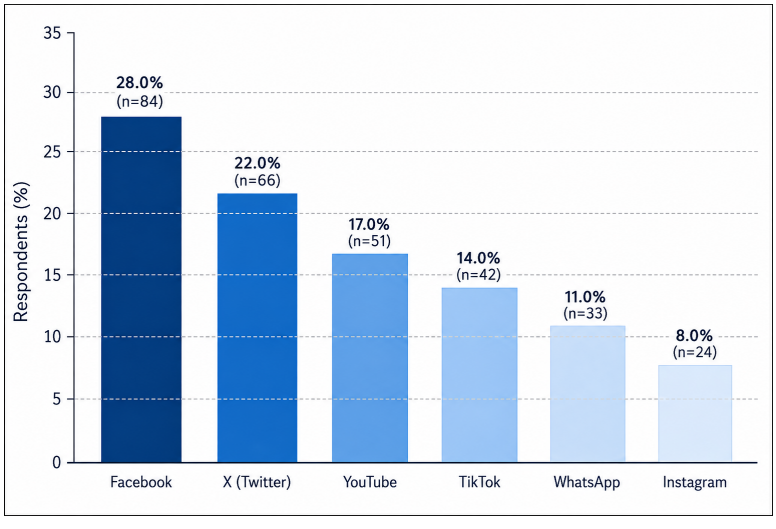
Variable	Category	Frequency (%)
Gender	Male	144 (48.0 %)
	Female	156 (52.0 %)
Age	18–24	78 (26.0 %)
	25–34	90 (30.0 %)
	35–44	66 (22.0 %)
	45–54	42 (14.0 %)
	55+	24 (8.0 %)
Education	Secondary	45 (15.0 %)
	Bachelor	132 (44.0 %)
	Master	96 (32.0 %)
	Doctorate	27 (9.0 %)
Daily platform use	< 1 hour	36 (12.0 %)
	1–3 hours	129 (43.0 %)
	3–5 hours	93 (31.0 %)
	> 5 hours	42 (14.0 %)

Encounters with Climate Misinformation

Participants indicated where they were most likely to encounter misinformation about climate change. Facebook led with 28.0 %, followed by X at 22.0 % and YouTube at 17.0 % (Figure 1). The findings suggest that the two key channels through which climate misinformation is diffused are visual platforms and algorithmically-driven platforms, in alignment with the diffusion patterns found in Vosoughi et al. (2018).

Figure 1

Platforms where Climate-change Misinformation is most Encountered (n=300)

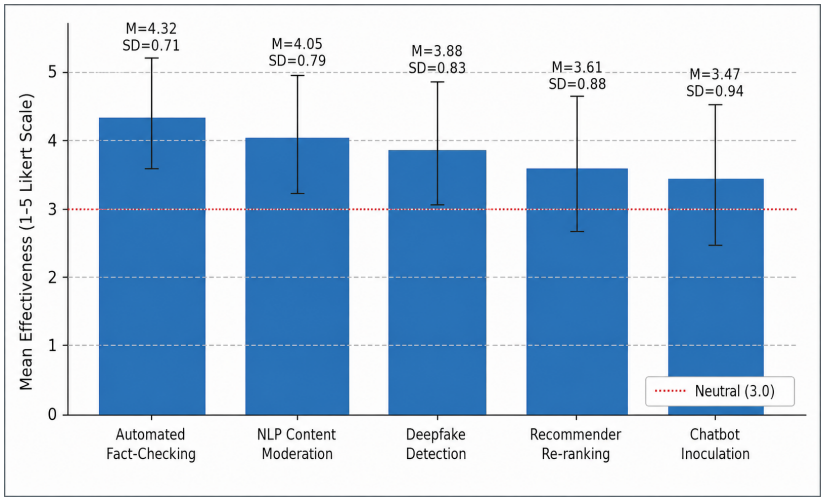


Perceived Effectiveness of AI Interventions

The mean effectiveness scores for five AI interventions are presented in Figure 2. Automated fact-checking received the highest rating (M = 4.32, SD = 0.71), followed by NLP content moderation (M = 4.05, SD = 0.79) and deepfake detection (M = 3.88, SD = 0.83). The overall findings suggest positive user experiences with AI-driven interventions, with all five tools showing ratings above the neutral score of 3.0.

Figure 2

Perceived Effectiveness of AI-driven Interventions (n=300)



Change in Weekly Misinformation Exposure

To determine if interactions with the AI-driven interventions reduced weekly exposure to climate misinformation, we used a paired-sample t-test. As shown in Table 2, mean weekly encounters declined significantly from 9.24 to 5.12 ($t = 12.87$, $p < .001$, Cohen's $d = 0.74$), representing a 44.6 % reduction. This drop is shown by age in the figure below.

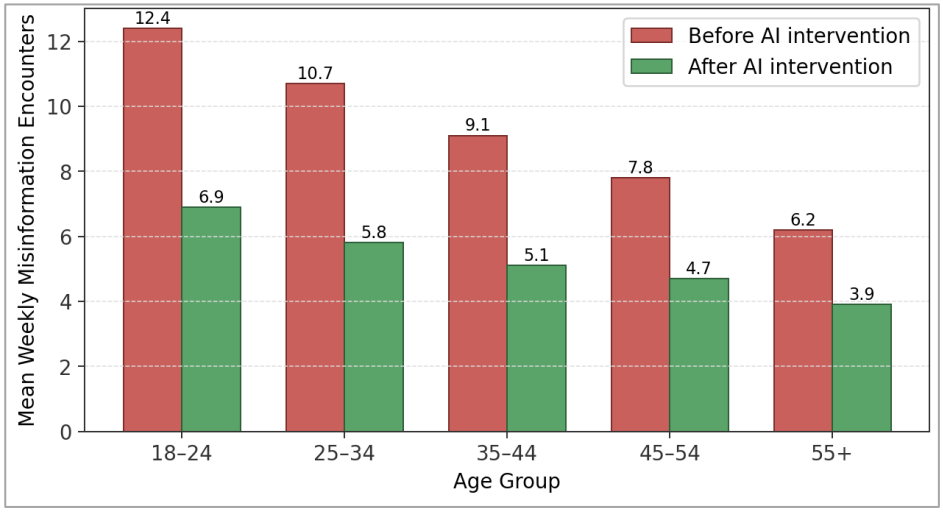
Table 2

Paired-Sample t-Test for Weekly Misinformation Exposure (N = 300)

Condition	Mean	SD	t	p
Before AI intervention	9.24	3.86		
After AI intervention	5.12	2.71	12.87	< .001

Figure 3

Change in Weekly Misinformation Encounters by Age Group (n=300)



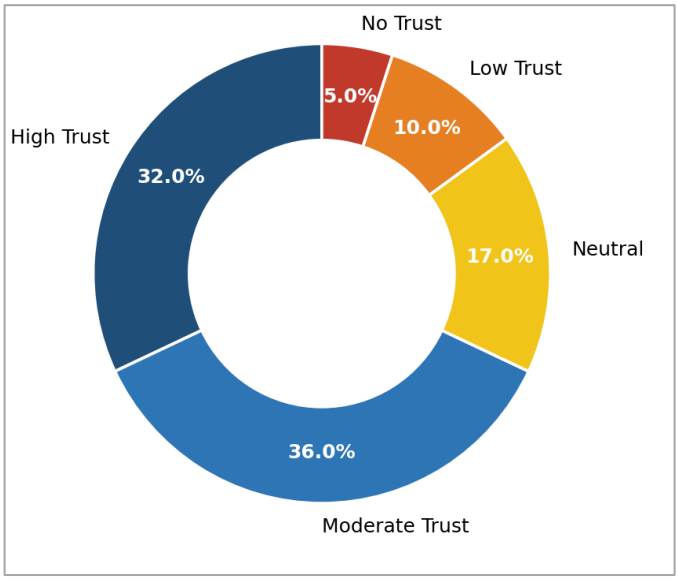
As seen in Figure 3, while younger respondents (those aged 18-24) were more likely to be exposed to misinformation than older respondents, they also seem to have experienced the greatest absolute drop, indicating that AI moderation has a significant effect on high-frequency digital users.

Trust in AI-Verified Climate Information

The distribution of trust in AI-verified climate information is reported in Figure 4. Regarding AI-verified content, 68% of respondents rated it as moderately or highly trustworthy, and 15% indicated that they have little or no trust. The pattern suggests a positive overall attitude toward the use of AI for mediation, but there is a significant group expressing a neutral stance (17 %), reflecting a level of scepticism regarding the transparency of algorithms.

Figure 4

Respondent Trust in AI-verified Climate Information (n=300)



Group Differences (One-Way ANOVA)

To test the differences in perceived effectiveness between age groups, a one-way ANOVA test was conducted. Findings showed that the perceived effectiveness of AI interventions was statistically significant, $F(4, 295) = 6.42$, $p < .001$, $\eta 2 = 0.08$, with age being the significant predictor. A post hoc Tukey test indicated that the mean difference between age range 25-34 and age range 55+ in the digitally native age group was significant at 0.61 ($p = .002$).

Table 3

One-Way ANOVA: Perceived Effectiveness by Age Group

Source	SS	df	F	p
Between groups	15.42	4	6.42	< .001
Within groups	177.05	295		
Total	192.47	299		

In general, the findings indicate that AI interventions can reduce exposure and engagement with climate misinformation significantly, but with demographic and platform disparities.

Discussion

The findings offer strong evidence that AI-powered interventions have a real impact on users' experience and understanding of climate change misinformation. This finding of automated fact-checking being considered the most effective tool (M = 4.32) mirrors the increasing capabilities of machine-learning classifiers as described by Coan et al. (2021), who developed a high-performing transformer-based model for detecting climate claims that were

counter to the general consensus. Likewise, the high rating of NLP content moderation (M = 4.05) aligns with the taxonomy proposed by Zhou and Zafarani (2020), in which they noted the flexibility of NLP pipelines for real-time detection of misinformation.

The weekly misinformation encounters after weeks of exposure to AI interventions are reduced by 44.6 % — a pattern similar to that observed by Pennycook and Rand (2019), who reported that algorithmic signals in combination with user reasoning significantly reduced belief in false content. The significant increase for younger participants matches findings by Diehl et al. (2021), who found that younger respondents are more likely to respond to corrections embedded in the platform, likely due to their increased exposure to such corrections.

Trust patterns are also informative. The 68 % of respondents who expressed moderate to high trust in AI-verified content suggests a cautious, yet positive, public mindset regarding this type of content, aligning with the inoculation effect of Roozenbeek and van der Linden (2018) and van der Linden et al. (2017), showing that exposure to weak misinformation creates long-term cognitive resistance. The 17 % neutral cluster, however, shows that there remains scepticism towards algorithmic opacity, similar to the concern expressed by Lewandowsky (2021) for the credibility of platform-level moderation.

Lastly, the large age-related differences are evidence that AI-driven climate interventions are not accepted on equal terms. This corresponds to the polarization trends observed by Falkenberg et al. (2022) in social media polarization regarding climate issues. Overall, the results indicate that AI holds great potential to address misinformation, but it must be designed, implemented, and engaged with publicly in culturally appropriate ways.

Research Implications

The results give concrete guidance to platform designers, policymakers and environmental communicators. Explainable AI fact-check systems, along with deepfake detection and inoculation chatbots in mainstream digital media, can substantially reduce the dissemination of climate misinformation. Explainable AI fact-checkers, deepfake detectors, and inoculation chatbots can be mainstreamed to measurably reduce the dissemination of climate misinformation. The report also emphasizes the need for algorithmic governance and specific measures for various target groups, particularly the elderly, to boost trust in AI-validated environmental information and promote informed climate decision-making.

Limitations of the Study

There are a few things to be aware of. One problem with the study is that the data were self-reported from the survey, which may be subject to social desirability bias and recall bias. Second, the geographic location of the sample of 300 respondents is concentrated and may not be generalisable across cultures. Third, the cross-sectional study measures perceptions at one point in time, and not long-term behaviour change. Fourth, the study concentrated on five AI-supported interventions and did not cover hybrid or human-in-the-loop moderation models. Lastly, generative AI technologies could quickly change the misinformation landscape and the efforts to respond to it, requiring replication and ongoing monitoring.

Conclusion and Recommendations

This research set out to delve into the potential role of artificial intelligence in combating climate change misinformation in digital media. The study, which was developed from the recommendations of 300 active users after an extensive analysis of academic literature, shows that these AI tools can significantly reduce false climate information per week, especially automated fact-checking, NLP-based content moderation, and detecting deep fakes.

The findings of a 44.6% decrease in encounters with misinformation and 68 % of the respondents feeling moderately or highly trusting of the AI-verified content provide strong empirical evidence of the growing role of AI in climate communication ecosystems. Nevertheless, their scepticism in certain segments of society, particularly the elderly, indicates that technological capacity is insufficient, and that it needs to be complemented by clear governance, media literacy and cross-cultural adaptation. Based on these results, the following recommendations are made. Firstly, digital media platforms need to adopt explainable AI systems that give reasons for flagging, downranking, or contextualising a specific climate post, reducing concerns of opaque algorithms. Secondly, policies should be developed to regularly evaluate AI moderation tools and implement the science-based consensus while acknowledging valid public discourse. Thirdly, environmental communicators and educators should integrate AI-based inoculation chatbots and fact-checking interactive assistants in the field of education on climate change, especially for high-risk populations such as adolescents and the elderly. Fourth, partnerships and collaboration between researchers, tech companies, and civil society across platforms of technologies should be created to share datasets, model benchmarks, as well as best practices in identifying emerging trends in climate misinformation, including deepfakes using AI. To conclude, future research could be conducted by taking a longitudinal or mixed methods approach to examine changes in behavior over time and to determine the impact of AI moderation on human cognition. When combined, these can help create the credibility of information, trust in the public by climate scientists and help coordinate action for climate change in the 21st century on a global scale in digital media.

References

Coan, T. G., Boussalis, C., Cook, J., & Nanko, M. O. (2021). Computer-assisted classification of contrarian claims about climate change. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-01714-4

Cook, J., Lewandowsky, S., & Ecker, U. K. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLOS ONE*, *12*(5), e0175799. https://doi.org/10.1371/journal.pone.0175799

Diehl, T., Huber, B., Gil de Zúñiga, H., & Liu, J. (2021). Social media and beliefs about climate change: A cross-national analysis of news use, political ideology, and trust in science. *International Journal of Public Opinion Research*, *33*(2), 197-213. https://doi.org/10.1093/ijpor/edz040

Falkenberg, M., Galeazzi, A., Torricelli, M., Di Marco, N., Larosa, F., Sas, M., Mekacher, A., Pearce, W., Zollo, F., Quattrociocchi, W., & Baronchelli, A. (2022). Growing polarization around climate change on social media. *Nature Climate Change*, *12*(12), 1114-1121. https://doi.org/10.1038/s41558-022-01527-x

Farrell, J., McConnell, K., & Brulle, R. (2019). Evidence-based strategies to combat scientific misinformation. *Nature Climate Change*, *9*(3), 191-195. https://doi.org/10.1038/s41558-018-0368-6

Lewandowsky, S. (2021). Climate change disinformation and how to combat it. *Annual Review of Public Health*, *42*(1), 1-21. https://doi.org/10.1146/annurev-publhealth-090419-102409

Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, *116*(7), 2521-2526. https://doi.org/10.1073/pnas.1806781116

Roozenbeek, J., & Van der Linden, S. (2018). The fake news game: Actively inoculating against the risk of misinformation. *Journal of Risk Research*, *22*(5), 570-580. https://doi.org/10.1080/13669877.2018.1443491

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06930-7

Treen, K. M., Williams, H. T., & O'Neill, S. J. (2020). Online misinformation about climate change. *WIREs Climate Change*, *11*(5). https://doi.org/10.1002/wcc.665

Ullah, A., Islam, K., Ali, A., & Baber, M. (2024). Assessing the impact of social media addiction on reading patterns: A study of Riphah International University students. *International Journal of Human and Society*, *4*(1), 1250-1262.

Ullah, A. (2026). Social media integration for modern library and information service promotion. *Journal of Digital Scholarship in Archives and Information*, *2*(1), 37-51. https://doi.org/10.58723/jdsai.v2i1.172

Van der Linden, S., Leiserowitz, A., Rosenthal, S., & Maibach, E. (2017). Inoculating the public against misinformation about climate change. *Global Challenges*, *1*(2). https://doi.org/10.1002/gch2.201600008

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

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