

# Spread of Misinformation in Malayalam: A Case Study on the User Engagement and Impact of Third-party Fact-checking on Facebook

Communication &  
Journalism Research  
Jul - Dec 2023 | 12 (2) pp 1-25  
©The Author(s) 2023  
ISSN: 2348 – 5663  
Website: [www.cjrjournal.in](http://www.cjrjournal.in)  
DOI:10.5281/zenodo.10537426

**Habeeb Rahman Y P<sup>1</sup>**

Research Scholar, Department of Journalism and Mass Communication, University of Calicut

**Muhammadali Nelliullathil**

Associate Professor, Department of Journalism and Mass Communication, University of Calicut

## Abstract

In today's ever-evolving digital landscape, where information dissemination shapes our perception of reality, the study of user engagement and the pivotal role played by fact-checking come to the forefront. This research focuses on the Malayalam-speaking community on Facebook to explore how users navigate and interact with misinformation, as well as the pivotal impact of third-party fact-checking within this specific linguistic and regional context.

The prevalence of misinformation, particularly in socio-political contexts, serves as a catalyst for this investigation. Kerala's distinguished political literacy and varying levels of digital proficiency add complexity to the misinformation landscape. Facebook's prominence in spreading misinformation emphasises the necessity for a thorough examination in the Malayalam speaking community.

By delving into the ways individuals in Kerala engage with misinformation on Facebook, this study casts light on their encounters with deceptive content. Additionally, it underscores the indispensable role played by third-party independent fact-checking initiatives in countering misinformation. The findings from this research would transcend academic boundaries, offering valuable insights for formulating effective strategies and policies aimed at upholding the integrity of public discourse and information accuracy within the digital realm.

**Keywords:** *Misinformation, Facebook, User engagement, Third-party fact-checking*

---

<sup>1</sup> Correspondence: [habeebrahman@uoc.ac.in](mailto:habeebrahman@uoc.ac.in)

## **Introduction**

In the digital era, misinformation has emerged as a critical social concern, exerting its influence through the vast and interconnected network of social media platforms. Among these platforms, Facebook holds a prominent place, serving as a magnifier of misinformation. The widespread reach and user engagement facilitated by Facebook amplify the impact of false information disseminated through this medium.

Within this digital landscape, it is imperative to examine the specific context of misinformation in the Malayalam language, as it holds cultural and linguistic significance, primarily in the Indian state of Kerala and among the global Malayali diaspora. Understanding the characteristics of misinformation within this linguistic framework is essential for formulating targeted strategies to combat its adverse effects. Misinformation manifests in diverse categories, each with its distinct consequences. These categories encompass political misinformation, communal discord, public health-related falsehoods, and more. Categorising misinformation aids in devising effective countermeasures.

The repercussions of misinformation are far-reaching, affecting public opinion, social cohesion, and even public health. The dissemination of false information can lead to confusion, polarise communities, and in some cases, pose a direct threat to individuals' well-being. Recognizing the gravity of these consequences underscores the urgency of addressing misinformation. The response of users to misinformation on social media is a pivotal aspect of this dynamic. User engagement with false information, as expressed through interactions such as likes, shares, and comments, plays a pivotal role in amplifying or mitigating the impact of misinformation. User responses, therefore, hold significant potential for shaping the narrative.

In combating misinformation, fact-checking serves as a crucial line of defense. Fact-checking organisations and initiatives scrutinise the veracity of claims and debunk false information. Their role in verifying and rectifying inaccuracies contributes to the credibility of information in the digital space. Third-party fact-checking initiatives operating on Facebook have gained prominence in recent years. These initiatives collaborate with the platform to assess and label potentially false information. Understanding the scope and effectiveness of third-party fact-checking on Facebook is vital for evaluating the platform's efforts to curtail the spread of misinformation.

As these platforms continue to serve as primary sources of information for a global user base, the need to address misinformation at its source becomes increasingly critical. Mitigating misinformation is not only a matter of digital literacy but a fundamental pillar of preserving the integrity of public discourse in the digital age.

## **Significance of the Study**

The research carries substantial significance in several key aspects. Firstly, it provides invaluable insights into the extent and nature of misinformation propagation on Facebook within the specific context of Kerala, illuminating the platform's role in disseminating different categories of false information. Additionally, the study facilitates a comprehensive evaluation of the effectiveness of third-party fact-checking systems within this linguistic community, shedding light on the mechanics and impact of these initiatives.

In a broader context, this study extends its importance by highlighting the importance for media literacy efforts beyond fact-checking. It underscores the need for a sophisticated approach to tackle misinformation, emphasising the significance of educating users to critically assess information and fostering a discerning audience.

## **Review of Literature**

In the landscape of combating misinformation on social media, Buchanan (2020) explored factors influencing the spread of disinformation, revealing that the authoritativeness of the source and consensus indicators had limited impact. Digital literacy showed little influence, highlighting the significance of pre-existing attitudes and familiarity with materials in predicting the likelihood of sharing disinformation. This emphasises the need for nuanced strategies to counter disinformation effectively.

According to Epstein et al. (2022), social cues influence users' engagement with social media posts. The presence of social cues, indicating a larger audience engagement, not only increased content sharing but also enhanced truth discernment. This unexpected finding suggests the complex dynamics of social cues and their impact on the dissemination of misinformation, urging platforms to carefully consider their design decisions.

Glenski et al. (2020) shifted the focus to understanding audience engagement with digital deception. By addressing key research questions about who engages with misinformation, how quickly they engage, and what feedback they provide, the study contributes valuable insights. Recognizing patterns and behaviours can inform strategies to enhance media literacy and encourage informed engagement with crowd-sourced information on social news platforms.

There is a decline in interactions with false content on Facebook while an increase persisted on Twitter. This shift in engagement patterns underscores the evolving

nature of the misinformation problem on different social media platforms, urging a nuanced approach to address platform-specific challenges (Allcott et al., 2019).

Perceived information characteristics, self-expression, and socialising motivations as significant factors prompting regular social media users to share misinformation. Considering gender differences, women report a higher prevalence of misinformation sharing. It highlights the multifaceted nature of motivations behind misinformation sharing (Chen et al., 2015).

Pal and Loke (2019) examined the features of fact-checking websites, contributing to the development of a framework for understanding their information work dimensions. This research provides practical implications for designing effective fact-checking websites and emphasises the role of these platforms in countering misinformation.

According to McDougall's (2019) media literacy education is a proactive approach to build resilience against 'information disorder' which aligns with the evolving challenges in the media landscape. Critical media literacy, when integrated into education, outperforms reactive resources, emphasising the need for comprehensive educational strategies to equip individuals with the skills to navigate an information-rich environment.

While acknowledging positive outcomes of media literacy initiatives, there are challenges such as a lack of comprehensive evaluation data, potential ineffectiveness for specific materials, and the risk of fostering overconfidence. The recommendations put forward by Bulger and Davison (2018) underscore the perpetual struggle in assessing and updating media literacy programs.

## **Operational Definitions**

### ***Misinformation***

Misinformation, in the context of this study, pertains to the dissemination of incorrect or false information, whether intentional or unintentional, with a focus on information that has the potential to mislead or compromise its accuracy. It encompasses a broad spectrum of inaccuracies, including but not limited to disinformation (intentionally false information) and mal-information (information shared with a negative intent). In this paper, the term "misinformation" encompasses all forms of compromised or erroneous information, acknowledging the diverse ways in which false information can be conveyed without making a specific distinction between various categories.

***Political misinformation:***

In the context of this study, "Political misinformation" refers to a specific category of misinformation characterised by content related to political matters. This category encompasses information, whether intentionally false or not, that is primarily designed to influence, promote, or defame a political party, personality, or ideology. Political misinformation often takes the form of political propaganda, where the dissemination of compromised information is used as a means to advance particular political agendas, shape public opinion, or discredit political entities. This operational definition emphasises the significant role of political content in the context of misinformation within the study's specified region.

***Health misinformation:***

In the context of this study, "Health misinformation" pertains to a distinct category of misinformation characterised by the dissemination of misleading information that contains claims related to health and wellness. This type of misinformation has the potential to mislead individuals regarding their health decisions, such as whether to seek treatment, take specific medications, or receive vaccinations. Health misinformation often poses a direct risk to human well-being by potentially causing harm or influencing individuals in ways that could be detrimental to their health and safety. The focus of this operational definition underscores the critical nature of health-related misinformation, particularly in terms of its potential to impact individuals' decisions and well-being within the context of health concerns.

***Third-party independent fact-checking:***

In the context of this study, "Third-party independent fact-checking" pertains to a process wherein fact-checkers, external to the platform, assess the accuracy of information disseminated on Meta platforms, including Facebook. Fact-checkers utilise their own investigative and reporting methods, which may involve interviewing primary sources, consulting publicly available data, and conducting thorough analyses of media content, such as photos and videos. This assessment is independent of content removal decisions, which are based on Facebook's community standards. When a piece of content is deemed false by a fact-checker, its distribution on the platform is notably reduced. Those who have previously shared or intend to share such content are notified that the information is false. A warning label, linking to the fact-checker's article disproving the claim with original reporting, is applied to the content.

## **Research Objectives**

This research is dedicated to understanding how users engage with and how third-party fact-checking impacts the spread of misinformation on Facebook, focusing on the Malayalam language. To achieve this, the study outlines a set of specific objectives designed to investigate these aspects within the Malayalam-speaking Facebook community.

- To analyse the misinformation ecosystem on Facebook in the context of Kerala
- To investigate user engagement with misinformation shared on Facebook
- To examine the impact of third-party fact-checking on Facebook

## **Methodology**

The methodology for this study primarily relies on qualitative content analysis. This approach is particularly well-suited for investigating the complex dynamics of misinformation, user engagement, and the influence of third-party fact-checking on Facebook within the Malayalam language context. Qualitative content analysis involves a systematic examination of the content available on the platform, allowing for an in-depth exploration of the themes, nuances, and patterns present in posts, comments, and fact-checking efforts.

The sample selection process is designed to capture a representative cross-section of fact-checks and their associated claims. To achieve this, a specific number of fact-checks and their corresponding claims were selected from a list of published fact-checks sourced from IFCN signatory fact-checking portals in Malayalam. The selected time frame for this sample spans from August 2023 to October 2023. This systematic approach to sample selection ensures a focused and comprehensive analysis of the information landscape during the specified period, providing insights into the prevailing misinformation and the fact-checking initiatives specific to the Malayalam-speaking Facebook community.

Due to the inherent constraints imposed by the qualitative nature of this study and the imperative to manage time and resource limitations, a select number of samples have been chosen for analysis. Specifically, the study focuses on fact-check articles produced by four IFCN-recognised Malayalam fact-checking portals, namely NewsMeter, India Today, News Checker, and Fact Crescendo, all of which serve as

independent third-party fact-checkers collaborating with Meta. Articles were selected through a purposive sampling approach, with each organisation contributing articles that were aligned with the study period, based on the findings of the previous research.

## Analysis

This chapter analyses the data collected from the digital realm, specifically Facebook. This comprehensive analysis examines various aspects, commencing with the content analysis of user engagement. We also evaluate the fact-checking process and its influence within the context of misinformation in the Malayalam language on Facebook. The analysis aims to explore the complex dynamics of misinformation, the role of third-party fact-checkers, and the resulting impact on the information landscape. This chapter provides valuable insights into how users, misinformation, and fact-checking efforts interact in the digital environment, offering a clearer understanding of this dynamic relationship.

### Primary analysis & categorization of sample data

Sl No	Claim	Posted by (As per FC article reference)	Posted Date	Fact-checked by/ on	Rating	Genre
1	Visuals of spacecraft landing on Mars	<a href="#">Unverified Facebook User</a>	26 June 2023	<a href="#">NewsMeter / 25 Aug 2023</a>	False	Uncategorized
2	Old photo during the election campaign of Oommen Chandy in Puthuppally constituency	<a href="#">Unverified Facebook Page</a>	21 Aug 2023	<a href="#">News Checker/ 24 Aug 2023</a>	False	Political

3	ISRO chairman visits RSS office	<a href="#">Unverified Facebook Page (Reel)</a>	27 Aug 2023	<a href="#">India Today/ 28 Aug 2023</a>	Missing Context	Political
4	Sangh Parivar attack on boy entered temple to drink water	<a href="#">Unverified Facebook User</a>	30 July 2023	<a href="#">Fact Crescendo/ 1 Aug 2023</a>	False	Communal
5	Islamic terrorists attacked army personnel in Kollam and wrote PFI on his back	<a href="#">Unverified Facebook Page</a>	25 Sep 2023	<a href="#">NewsMeter/ 27 Sep 2023</a>	False	Communal
6	Photoshoot by PM Narendra Modi at Ramoji Film City	<a href="#">Unverified Facebook User</a> <a href="#">(Archive Link)</a>	22 Sep 2023	<a href="#">News Checker/23 Sep 2023</a>	False	Political
7	Muslim boys and Hindu girls were caught from a cafe in MP during Police raid	<a href="#">Unverified Facebook User</a>	16 Sep 2023	<a href="#">India Today/22 Sep 2023</a>	Missing Context	Communal



8	The source of Nipah virus has been identified as Arecanut	<a href="#">Unverified Facebook User</a>	17 Sep 2023	<a href="#">Fact Crescendo/ 22 Sep 2023</a>	False	Health
9	Photo of father and daughters with a smile in a collapsed building at Palestine	<a href="#">Unverified Facebook User</a>	26 Oct 2023	<a href="#">NewsMeter / 26 Oct 2023</a>	False (AI Generated)	Political / Communal
10	Samsung offers mobile phones as gift for lucky draw through social media	<a href="#">Unverified Facebook Page</a>	8 Oct 2023	<a href="#">News Checker / 11 Oct 2023</a>	False	Scam
11	Union government announces maternity benefits for MGNREGA workers	<a href="#">Unverified Facebook page</a>	15 Oct 2023	<a href="#">India Today / 18 Oct 2023</a>	False	Political
12	Woman without Burqa was not allowed to travel in bus in Kasaragod, Kerala	<a href="#">Unverified Facebook Page</a>	27 Oct 2023	<a href="#">Fact Crescendo/ 27 Oct 2023</a>	False	Communal

Table 1.1 Sample data set

## **Misinformation ecosystem on Facebook**

An analysis of the sample data clearly shows that the majority of misinformation during the study period falls into the categories of politics and communal affairs. Kerala is well-known for its political awareness and harmonious coexistence of different religious and communal groups. Given this context, it becomes evident that a significant amount of misinformation, particularly within these two categories, reflects deliberate attempts to influence people politically or incite communal tensions through social media.

These initial observations are evident from the primary analysis of the collected data. The prevalence of misinformation in politics and communal matters raises questions about potential motives behind these misleading narratives. It suggests that misinformation campaigns might be carefully coordinated to politically manipulate individuals or sow discord within communities. The juxtaposition of political awareness and communal harmony with the prevalence of misinformation underscores the critical need to investigate the origins and motivations behind these deceptive narratives, which have the potential to disrupt the social and political fabric of the region.

While health misinformation is relatively less prevalent, it is contextually significant, especially during health crises. For instance, the data includes examples shared during the Nipah virus outbreak in the state. It highlights the contextual and potentially harmful nature of health misinformation. They emphasise the vital importance of promoting health literacy and implementing effective strategies to combat health-related misinformation.

Another noteworthy discovery from the initial data analysis is the prevalence of AI-generated content and scams, which pose significant threats. These types of misinformation can be particularly challenging for the general population to identify, potentially leading to more severe consequences. Scams, in particular, exploit individuals who may not possess strong digital literacy, making them vulnerable to financial fraud and similar risks. These findings underscore the need for heightened awareness and concern regarding these emerging issues.

In the evaluation of misinformation, a spectrum of ratings is observed. The majority of the selected samples consists of completely false claims lacking factual accuracy. Following closely are contents presented out of context, where the information itself is true but framed out of context. Partly false information is also present but in relatively fewer instances. Notably, there is the presence of edited or altered media contents, particularly in images and videos, including those generated by AI. This

primary analysis underscores a prevalent pattern that is the predominant circulation of entirely false information across the internet, indicating the potential for significant harm and consequences.

Identifying the origins of this misinformation is often a complex task, primarily due to technical constraints. However, it is important to note that the initial sharing of political and communal misinformation predominantly occurs through various Facebook pages, while other categories are often associated with individual user accounts. Also, a considerable portion of these accounts appears to be inauthentic, possibly representing fake profiles.

### **User engagement**

In exploring how users engage with misinformation on Facebook in the Malayalam-speaking community, the analysis uncovers diverse patterns. Initial scrutiny of cumulative comments on misinformation reveals a prevalent lack of awareness among users about the claims' accuracy. Many users respond with unwavering belief, contributing to the challenge of misinformation spread. The investigation extends to specific misinformation categories, highlighting variations in user responses. Wrong claims related to science and technology often receive more support from users, suggesting an increased vulnerability to misinformation. Political claims, in contrast, attract a discerning audience with commendable political literacy, although claims disseminated through political entities' pages present a paradox due to algorithmic effects and echo chambers. Communal claims, influenced by mainstream media, often see greater user recognition as misinformation. This section unravels the complexities of user engagement, revealing varying degrees of susceptibility and resilience across different thematic categories.

SI No	Claim	No, of Comments	Recognizing	Not Recognizing	Neutral	%
		(Cumulative numbers from select posts)				
1	Visuals of spacecraft landing on Mars	358	22	272	64	75.9%

2	Old photo during the election campaign of Oommen Chandy in Puthuppally constituency	424	29	328	67	77.3%
3	ISRO chairman visits RSS office	189	17	149	23	78.8%
4	Sangh Parivar attack on boy entered temple to drink water	214	17	168	29	78.5%
5	Islamic terrorists attacked army personnel in Kollam and wrote wrote PFI on his back	623	265	313	45	50.2%
6	Photoshoot by PM Narendra Modi at Ramoji Film City	214	12	178	24	83.2%
7	Muslim boys and Hindu girls were caught from a cafe in MP during Police raid	186	28	132	26	70.9%

8	The source of Nipah virus has been identified as Arecanut	236	45	152	39	64.4%
9	Photo of father and daughters with a smile in a collapsed building at Palestine	64	8	39	17	60.9%
10	Samsung offers mobile phones as gift for lucky draw through social media	443	12	409	22	92.3%
11	Union government announces maternity benefits for MGNREGA workers	89	21	56	12	62.9%
12	Woman without Burqa was not allowed to travel in bus in Kasaragod, Kerala	378	207	149	20	39.4%

*Table 1.2 User engagement analysis*

Upon scrutinising the collective comments on selected posts for each claim, a prevailing lack of awareness among users regarding the veracity of the claims becomes apparent. The majority of users exhibit a tendency to unquestioningly believe in the truthfulness of the claim, evident in their comments. Notably, the number of users who take the time to comment represents a subset of those who

engage more extensively with the post, while a larger segment merely reacts to the content. This suggests a potentially higher number of individuals who may endorse the claim as true, surpassing the representation indicated by comment analysis alone. Across all categories of misinformation examined in this study, approximately 60% of the comments fail to recognize the factual inaccuracy of the shared content. Within the remaining 40%, half of the responses adopt a neutral stance, abstaining from taking a definitive position on the accuracy of the information.

Another observation pertains to claims related to science and technology, where a lack of awareness among users is even more pronounced compared to claims of a communal nature. This indicates that misinformation within the domain of science and technology may have a more significant potential to mislead individuals in the specific context. Over 75 percent of comments within this category exhibit an inability to identify the shared content as inaccurate or misleading.

Analysing the comments associated with the most shared political claims, a higher proportion of individuals demonstrate an ability to recognise these claims as misinformation compared to other categories, underscoring the commendable political literacy within Kerala. However, a notable pattern emerges when examining claims disseminated through Facebook pages of political parties, politicians, or users with a substantial follower base affiliated with a specific political party. In such instances, comments in support of the claims outweigh dissenting opinions. This phenomenon can be attributed to the algorithmic features and echo chamber effect inherent in social media platforms. Users aligned with a particular political party tend to support claims against opposing parties, even when aware of their inaccuracies. This phenomenon does not necessarily imply an inability to recognise misinformation; rather, it highlights a tendency to actively propagate disinformation. The distinction between misinformation recognition and deliberate disinformation dissemination warrants further investigation into the complexities of political discourse on social media platforms.

Analysing comments on claims categorised as communal reveals a notable trend wherein a larger number of individuals demonstrate an ability to recognise these claims as misinformation. This discernment is notably influenced by the significant role played by mainstream media, as evidenced in one of the communal claims chosen for this study. The claim, later debunked by mainstream media, saw the widespread sharing of news clippings and links by numerous users under the post, aiding in disseminating the truth. This suggests that, despite instances of political polarisation, a substantial majority of users actively challenge communal claims, transcending their political affiliations. This finding also underscores the harmonious coexistence

and fraternity within the state across religions, distinguishing it from trends observed in other regions. However, it is acknowledged that instances of communal intolerance persist on social media, primarily confined within echo chambers, as the majority often publicly refutes them, leveraging mainstream media support. This complex interplay highlights the dual dynamics of communal discourse on social media – internalised polarisation and the broader, public defence of communal harmony, often mediated by mainstream media interventions.

### **Impact of third-party independent fact-checking**

In combating the spread of misinformation on Facebook, a critical component of the platform's strategy involves collaboration with third-party organisations dedicated to fact-checking. These organisations operate jointly with Meta to scrutinise the accuracy of information circulating on the platform. The fact-checking process involves these organisations preparing and publishing detailed fact-check articles on their respective websites. Fact-checkers then utilise these articles' links to assess and rate the accuracy of posts flagged as potentially false or misleading.

Once a post is rated as false or lacking context, the post owner receives a notification, alerting them that the shared information is not deemed factually accurate. Simultaneously, individuals who shared the information also receive notifications about the fact-check. An interesting aspect of this process is the option given to the post owner to appeal the fact-check decision. This involves presenting evidence in support of the initially shared claims. The appeal undergoes further scrutiny by the fact-checker, who may choose to withdraw the rating or reject the appeal based on the provided evidence.

Importantly, the application of a fact-check rating has tangible consequences for the post's reach and visibility. Once a post receives a fact-check rating, its reach declines, and visibility across platforms is significantly reduced.

This section delves into an in-depth analysis of the impact of third-party independent fact-checking on Facebook.

<b>Sl No</b>	<b>Claim</b>	<b>Claim Date</b>	<b>FC on</b>	<b>Total Comments (Multiple posts)</b>	<b>Before FC</b>	<b>After FC</b>
1	Visuals of spacecraft landing on Mars	26 June 2023	25 Aug 2023	358	337	21
2	Old photo during the election campaign of Oommen Chandy in Puthuppally constituency	21 Aug 2023	24 Aug 2023	424	368	56
3	ISRO chairman visits RSS office	27 Aug 2023	28 Aug 2023	189	171	18
4	Sangh Parivar attack on boy entered temple to drink water	30 July 2023	1 Aug 2023	214	Post deleted by the user(s) after fact-check rating	
5	Islamic terrorists attacked army personnel in Kollam and wrote wrote PFI on his back	25 Sep 2023	27 Sep 2023	623	374	249



6	Photoshoot by PM Narendra Modi at Ramoji Film City	22 Sep 2023	23 Sep 2023	214	Post deleted by the user(s) after fact-check rating	
7	Muslim boys and Hindu girls were caught from a cafe in MP during Police raid	16 Sep 2023	22 Sep 2023	186	168	18
8	The source of Nipah virus has been identified as Arecanut	17 Sep 2023	22 Sep 2023	236	Post deleted by the user(s) after fact-check rating	
9	Photo of father and daughters with a smile in a collapsed building at Palestine	26 Oct 2023	26 Oct 2023	64	57	7
10	Samsung offers mobile phones as gift for lucky draw through social media	8 Oct 2023	11 Oct 2023	443	Post deleted by the user(s) after fact-check rating	

11	Union government announces maternity benefits for MGNREGA workers	15 Oct 2023	18 Oct 2023	89	64	25
12	Woman without Burqa was not allowed to travel in bus in Kasaragod, Kerala	27 Oct 2023	27 Oct 2023	378	236	142

*Table 1.3 Analysis of the impact of third-party fact-check*

In the majority of instances, fact-check stories are promptly published within two to three days of the initial posting of the content. The immediacy of this process contributes to an almost instantaneous decline in the post's circulation. The speed and thoroughness of this decline, however, depend on the efficiency of the fact-checking organisations and the number of organisations involved in scrutinising the content. It is evident from our analysis that most fact-checking organisations exert considerable effort to expedite the publication of fact-check articles. Notably, there are instances within the sample set where a fact-check is published on the very same day as the content's initial posting. Additionally, a comprehensive examination of the fact-check sample set reveals a concerted effort by fact-check organisations to maintain diversity in their fact-checks, fostering independence in their evaluations. This commitment to diverse and timely fact-checking practices forms a crucial aspect of the overall impact and efficacy of fact-checking initiatives on Facebook.

The decline in the number of comments post the publication of fact-check stories and the subsequent rating of content by fact-checkers is a notable trend. However, it is crucial to interpret this reduction as a technical consequence of the Facebook platform's back-end processes. When a content piece undergoes fact-checking and receives a rating, its reach diminishes, naturally leading to a decrease in the number of comments.

Despite this general pattern, specific instances arise where the comments persist even after fact-checking. This phenomenon is particularly observable in posts where users

actively identify misinformation, presenting evidence within their comments. Such instances often coincide with media reports from mainstream regional outlets. This nuanced observation suggests that users tend to rely more on regional news media platforms than formal fact-check articles, especially when misinformation gains coverage in local news. In contrast, for the majority of cases, comments appearing after the publication of fact-check stories constitute approximately 10 percent of the total number of comments.

In certain instances, users opted to delete their content following the publication of fact-checks and rating of the posts. This action can be attributed to two potential motivations. Firstly, it might indicate users' awareness of the consequences associated with sharing false information, reflecting a positive inclination towards rectifying their online contributions. Alternatively, users might delete content strategically to enhance the visibility of their profiles or pages as keeping rated content could diminish reach and impact monetisation. Despite these instances, it is crucial to highlight that over 70 percent of the sampled claims persist on Facebook. This suggests that a substantial proportion of users are unwilling to retract or rectify their content, portraying a concerning indifference to the consequences even after being notified by the fact-check.

An additional aspect to consider is the language barrier that could impede users' comprehension of the fact-checking process. Notifications and labels associated with fact-checking are provided in English, which might pose a challenge for local users who may not fully grasp the details of the fact-checking procedure. While the fact-check articles themselves are presented in Malayalam, crucial instructions, warnings, and labels are not available in the regional language. This linguistic gap presents a potential hindrance, preventing users from fully understanding the truth unveiled by fact-checking efforts. Addressing this language disparity is pivotal in ensuring that users can access and comprehend the corrective information provided through fact-checks, enhancing the efficacy of these interventions in combating misinformation.

Upon analysing the user engagement and impact of third-party fact-checking, it is found that the decline in comments post fact-checking is largely technical, driven by the platform's back-end processes reducing post visibility. Notably, posts with user-identified misinformation sustain discussion, emphasising reliance on regional news media. The fact-checking process, initiated promptly within two to three days in most cases, impacts post circulation, showcasing the dedication of fact-check organisations. However, users' reluctance to withdraw or correct content, despite notifications, signals a concerning indifference to misinformation consequences. Moreover, the language barrier in fact-check notifications poses a challenge,

potentially hindering comprehension among local users. These findings underscore the need for targeted interventions to enhance user awareness, improve comprehension, and foster a collective commitment to combating misinformation within this unique socio-linguistic landscape.

## **Findings**

Navigating the complexities of user engagement and third-party fact-checking on Facebook within the Malayalam-speaking community, the analysis reveals various insights. As the study explores a number of findings, we uncover essential patterns in how users respond to fact-checking and how misinformation spreads. The following points distil these insights, covering aspects such as the technicalities of post visibility, user reactions to fact-check notifications, and the vital role of language comprehension in the effectiveness of fact-checking efforts.

- The visibility of posts experiences an immediate decline following fact-checking, primarily due to Facebook's back-end processes.
- A significant number of users exhibit difficulty in identifying misinformation, as evident from the analysis of comments.
- Posts containing user-identified misinformation tend to sustain discussion, indicating a reliance on regional news media for validation.
- Fact-checks are typically published within two to three days of the content posting, underscoring the dedication and efficiency of fact-checking organisations.
- Users exhibit a reluctance to withdraw or correct content even after receiving fact-check notifications, reflecting a concerning indifference to the consequences of misinformation.
- The language barrier in fact-check notifications, with crucial instructions, warnings, and labels in English, presents a challenge for local users, potentially hindering full comprehension.

## **Discussion and Recommendations**

The effectiveness of third-party fact-checking, introduced by Meta to curtail the spread of false information on its social media platforms, is evident. This strategic approach adeptly achieves its technical objective of diminishing the reach and circulation of misleading content. However, challenges arise in effective communication and user education due to language barriers, necessitating adaptations

to enhance accessibility. To address this, it is crucial to make fact-checking processes more regional language-friendly, fostering better understanding among users. Moreover, platforms like Facebook should lead the development of user-friendly tools and techniques empowering individuals to verify information before posting. Urgent advancements in these areas are important, considering the complex nature of online misinformation. This proactive approach is essential for cultivating a more informed and vigilant digital community.

Analysing user comments reveals a prevalent inability among a substantial portion to recognise misinformation, evident from their comments devoid of factual understanding. This underscores the imperative to educate users on identifying misinformation through media literacy. While a majority remains unaware, it's crucial to acknowledge that some users actively support false claims for political reasons. This highlights the need to impart knowledge about the social responsibility of combating misinformation. Initiatives focused on media and information literacy can significantly contribute to enhancing users' abilities to navigate the digital landscape responsibly.

The role of regional media emerges prominently in addressing false information spread on social media platforms. When regional media outlets pick up misinformation issues, amplify them through news coverage, and share on their social media handles, the wider visibility lends credibility to the fact that the information is false. Consequently, a majority of users who inadvertently shared the misinformation often correct themselves, referencing the credible media posts. Interestingly, this corrective mechanism is less pronounced with fact-check stories, as users seldom engage by commenting with links to fact-checks. This underscores the substantial impact and responsibility that legacy media platforms hold in the battle against misinformation. Their vigilance in fact-checking, particularly through social media channels, can significantly contribute to cultivating a more accurate and informed digital discourse.

While fact-check organisations demonstrate commendable effectiveness in promptly publishing fact-check stories as soon as misinformation starts circulating, they encounter the challenge of limited readership. To maximise their impact, fact-checks should be disseminated through more innovative and interactive media platforms, reaching a broader audience. These organisations must strategically enhance their visibility and actively educate social media users about the essence and utility of fact-checking. Despite their substantial potential, the key lies in reaching out to a more extensive user base, ensuring that the valuable service they provide becomes an integral part of the broader digital discourse. Proactive outreach efforts can harness

the full potential of fact-check organisations in the ongoing battle against misinformation.

The reluctance of users to withdraw their posts even after fact-checking often stems from a lack of understanding about the notifications and their implications. This can be attributed to a combination of factors, including ignorance and a general unawareness of the potential consequences of sharing misinformation. Educating users on the nuances of misinformation, its repercussions, and fostering media literacy are crucial steps to empower individuals in navigating social media responsibly. Encouraging a mindset for correction, where users promptly rectify misinformation upon realisation, is essential. Language barriers can significantly contribute to this reluctance, emphasising the need for platforms like Meta to take proactive measures. Introducing notifications and instructions in regional languages would enhance accessibility and understanding, catering to the diverse linguistic landscape of users. Also, there is a greater possibility of pre-bunking to be carried out by various media organisations to make users aware about instances of misinformation even before they encounter them.

### **Limitations of the Study**

This study is subject to certain technical limitations. The sample set, derived through purposive sampling from all IFCN fact-check portals in Malayalam, may not comprehensively represent the entire misinformation ecosystem. The small sample size, a consequence of time constraints inherent in qualitative studies, potentially affects the generalisability of the identified misinformation landscape.

Moreover, some posts were deleted after fact-checks were published, introducing accessibility limitations for analysing comments and user engagement both pre and post fact-check. Additionally, as users have the autonomy to delete their content on Facebook and archiving is not feasible, referencing the study may encounter challenges, and figures may undergo changes after the research paper's publication.

Also, technical constraints impeded the retrieval of readership data for fact-check articles. Consequently, an accurate analysis of the reach of fact-check articles was not feasible, underscoring an additional limitation in assessing the broader impact of third-party fact-checking efforts.

Despite these limitations, it is essential to note that concerted efforts were made to provide accurate information, resulting in meaningful findings that contribute to our understanding of misinformation dynamics in the digital space.

## **Conclusion**

The study brings out the complex dynamics of misinformation within the Malayalam-speaking community on Facebook, emphasising the critical need for effective strategies in combating false information. Delving into the digital realm, this study has revealed various aspects of user engagement and the impact of third-party fact-checking.

The examination of user engagement patterns revealed a prevalent difficulty among users in discerning misinformation, necessitating media and information literacy initiatives. The findings also shed light on the effectiveness of third-party fact-checking in curtailing the circulation of false information on social media platforms. However, the limitations inherent in the study, such as sample representativeness and the challenges of analysing deleted posts, underscore the complex nature of researching misinformation in the digital age.

Beyond these limitations, the research serves as a crucial step toward understanding the contours of the misinformation landscape in the Malayalam-speaking community. The insights derived from the analysis of user engagement and the impact of third-party fact-checking contribute valuable perspectives to ongoing discourse on combating misinformation. This study reinforces the need for a multi-pronged approach, combining technological innovations, media literacy campaigns, and collaboration with fact-checking organisations to navigate the evolving challenges of misinformation on social media.

As digital landscapes continue to evolve, it is important to adapt strategies continually, considering the evolving nature of misinformation. The findings of this research pave the way for informed interventions, emphasising the collective responsibility of platforms, media organisations, and users in fostering an environment that upholds the integrity of information dissemination in the digital era.

## ***Conflict of Interest Declaration***

I declare that there is no conflict of interest regarding the research presented in my article. I am not associated with any organization that has a financial interest in the subject matter or the data/materials used in the article.

## References

- Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, 6, 2053168019848554. <https://doi.org/10.1177/2053168019848554>
- Buchanan, T. (2020). Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. *Plos One*, 15(10), e0239666. <https://doi.org/10.1371/journal.pone.0239666>
- Bulger, M., & Davison, P. (2018). The promises, challenges, and futures of media literacy. *Journal of Media Literacy Education*, 10(1), 1-21. <https://doi.org/10.23860/JMLE-2018-10-1-1>
- Carlos, D. (2023, September 22). Fake claim that the source of Nipah has been found as arecanut. Fact Crescendo. <https://www.malayalam.factcrescendo.com/false-propaganda-in-the-name-of-nipah-claims-nipah-virus-source-was-found-from-areca-nut-sample/>
- Chen, X., Sin, S.-C. J., Theng, Y.-L., & Lee, C. S. (2015). Why do social media users share misinformation? In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 111–114). <https://doi.org/10.1145/2756406.2756941>
- Epstein, Z., Lin, H., & Pennycook, G. (2022). How many others have shared this? Experimentally investigating the effects of social cues on engagement, misinformation, and unpredictability on social media. <https://doi.org/10.48550/arXiv.2207.07562>
- Glenski, M., Volkova, S., & Kumar, S. (2020). *User Engagement with Digital Deception*. [https://doi.org/10.1007/978-3-030-42699-6\\_3](https://doi.org/10.1007/978-3-030-42699-6_3)
- Laila, S. (2023, August 28). Is this the video of ISRO chairman visiting RSS office? India Today. <https://malayalam.indiatoday.in/fact-check/story/fact-check-the-video-of-the-isro-chairman-visiting-the-rss-office-is-misleading-659629-2023-08-28>
- Laila, S. (2023, October 18). Maternity benefits for MGNRES workers are provided by the Kerala government. India Today. <https://malayalam.indiatoday.in/fact-check/story/fact-check-mgnreg-workers-receive-maternity-benefits-under-a-kerala-government-scheme-696966-2023-10-18>
- McDougall, J. (2019). *Media literacy versus fake news*. Medijske studije.



- Pal, A., & Loke, C. (2019). Communicating fact to combat fake: Analysis of fact-checking websites. In Proceedings of the 2019 International Conference on Information Technology and Computer Communications (pp. 66–73). <https://doi.org/10.1145/3355402.3355415>
- Puzhakkal, D. (2023, September 22). Video from a cafe in UP spreads with false love jihad spin. India Today. <https://malayalam.indiatoday.in/fact-check/story/fact-check-viral-video-of-cops-nabbing-couples-during-raid-shared-with-false-love-jihad-spin-682461-2023-09-22>
- Rahman, H. (2023, August 25). Know the fact of video showing spacecraft landing on Mars. NewsMeter. <https://newsmeter.in/fact-check-malayalam/video-of-spacecraft-landing-on-mars-716972>
- Rahman, H. (2023, September 27). Who wrote PFI on the back of army personnel in Kollam? Know the truth. NewsMeter. <https://newsmeter.in/fact-check-malayalam/pfi-workers-attacked-indian-army-jawan-in-kerala-and-wrote-pfi-on-his-back-718470>
- Rahman, H. (2023, October 26). Know the truth behind father and daughters smile in the background of a collapsed building. NewsMeter. <https://newsmeter.in/fact-check-malayalam/picture-of-father-and-daughter-smiling-in-a-collapsed-building-at-palestine-719806>
- Thomas, S. (2023, August 24). The election campaign picture of Oommen Chandy is not from Puthupally. News Checker. <https://newschecker.in/ml/fact-check-ml/news-ml/not-of-puthupally/>
- Thomas, S. (2023, September 23). Is this the Photoshoot of PM Narendra Modi at Ramoji Film City? News Checker. <https://newschecker.in/ml/fact-check-ml/news-ml/not-modi-photoshoot/>
- Thomas, S. (2023, September 11). Fake lucky draw campaign on the name of Samsung on Social Media. News Checker. <https://newschecker.in/ml/fact-check-ml/viral-ml/fake-lucky-dip>
- Vasuki, S. (2023, August 1). Boy gets beaten by Sangh Parivar for entering the temple for drinking water. Fact Crescendo. <https://www.malayalam.factcrescendo.com/fact-check-the-image-of-boy-beaten-by-sangh-parivar-for-drinking-water-from-temple/>
- Vasuki, S. (2023, October 27). Did girls ask a woman to get out of the bus for not wearing a Burqa in Kasaragod? Fact Crescendo. <https://www.malayalam.factcrescendo.com/muslim-girl-students-verbally-abusing-a-hindu-woman-for-entering-the-bus-a-fact-check/>