J. DIFERENSIAL/E-ISSN:2775-9644
Vol.8(1),2026, 20–32
doi:https://doi.org/10.35508/jd.v8i1.26853

RESEARCH ARTICLE

Analysis of the COVID-19 Agenda Using Big Data from Social Media: A Comparative Study across Countries with R Programming

Şakir İşleyen[1], Amar Yahya Zebari[2,*], Hasan Hazim Jameel[3]

[1]*College of Economic and Administration, Department of Economic, Van Yuzuncu Yil University,Van, Türkiye*
[2]*Department of Business Administration, Cihan University-Duhok, Duhok, Iraq*
[3]*College of Education, Department of Mathematics, University of Duhok, Duhok, Iraq*
[*]*Corresponding author: Amar.yehya85@gmail.com*

Received: 26 December 2025; Revised: 18 January 2026; Accepted: 6 February 2026; Published: 4 March 2026.

Abstract:
Social media platforms are becoming increasingly important as sources of public discourse and real-time data analysis, as the COVID-19 epidemic has highlighted. Using the hashtag #COVID19, this study examines COVID-19-related tweets from seven nations (the US, Germany, South Korea, Iraq, Spain, Italy, and Türkiye) in order to find trends in engagement and correlations. Similarities between public attitude and government communications are examined by statistical techniques such as content analysis, frequency analysis, and cross-delay correlation, as well as R programming. The findings show that tweet patterns from different countries are highly correlated, and that the Iraqi government's tweets with a typical theme were more popular than those with a COVID-19 theme. This study provides information on cross-border communication tactics in times of crisis and illustrates the potential of big data analytics for comprehending global phenomena.

Keywords: Big Data, Social Media Analytics, COVID-19, R Programming, Cross-Country Analysis.

1. Introduction

The COVID-19 pandemic represents one of the most significant global crises of the 21st century, affecting every aspect of life, including health systems, economies, and societal behaviour [1]. The unprecedented spread of the virus and its impact have generated an overwhelming amount of information and discourse, much of which is shared on social media platforms. Among these platforms, Twitter has emerged as a critical tool for real-time information sharing, public engagement, and sentiment expression. Big data analytics provides an opportunity to process this voluminous and unstructured information, turning it into actionable insights [2]. Social media content, particularly tweets, offers a unique lens to study public perceptions, governmental communication strategies, and the overall societal response to the pandemic. Leveraging statistical tools and programming languages such as R, researchers can efficiently analyze large-scale social media data to uncover patterns, correlations, and temporal trends. The agenda-setting theory emphasizes the role of media in shaping

public attention and defining the issues perceived as socially salient, a process that has increasingly extended to digital and social media environments [3–5].

With the advent of Web 2.0, this theory has been extended to social media platforms, where user-generated content often influences traditional media narratives. Previous studies have employed cross-media analysis and sentiment analysis to explore these dynamics. However, few have focused on leveraging big data techniques to analyze pandemic-related communication on social media. The emergence of social media has prompted scholars to revisit agenda-setting theory within digitally networked communication systems. Meraz [6] demonstrates that agenda-setting dynamics in online environments are characterized by multidirectional influence, where traditional media, political actors, and online publics interact through time-sensitive information flows. Similarly, Conway et al. [7] provide empirical evidence of intermedia agenda-setting effects on Twitter, showing that the platform not only reflects political agendas but also contributes to their diffusion and amplification across media channels. These studies suggest that Twitter operates as both an agenda carrier and an agenda influencer, challenging hierarchical models of agenda-setting and emphasizing the importance of temporal and relational analyses in social media research.

The literature on agenda-setting and media influence reflects a fascinating evolution, from traditional frameworks to the dynamic interplay of digital platforms and public discourse. Wicke and Bolognesi [8] analyzed Covid-19 discourse on Twitter, examining topic shifts, sentiment changes, subjectivity, and figurative framing over time. Their findings highlight how public perception and engagement evolved during the pandemic, offering insights into social media communication dynamics in crisis contexts. Similarly, Mundt et al. [9] examined how social media facilitates the expansion and internal strengthening of social movements, focusing on Black Lives Matter. Their study highlights social media's role in building connections, mobilizing resources, forming coalitions, and amplifying alternative narratives, while also addressing associated challenges and risks.

Zhang and Zhu [10] investigated how source-level and message-level factors affect journalists' social media visibility during public health crises. Their study reveals that while branding has minimal impact, both affective and rational messages, as well as tweets containing images or news-related hyperlinks, receive higher engagement. Ittefaq et al. [11] conducted a computational analysis comparing COVID-19 news coverage across six major newspapers from China, South Korea, the United Kingdom, and the United States. The study identified both similarities and differences in the framing of the pandemic, noting that Western outlets were more likely to employ fear appeals in their reporting. These findings underscore the influence of journalism cultures on media coverage during global health crises. Earl et al. [12] examined the phenomenon of digital repression, analyzing how governments and other actors utilize digital tools to suppress social movements, protests, and activism. The research highlights various methods of digital repression, including surveillance, censorship, and disinformation campaigns, and discusses the implications for civil liberties and democratic engagement in the digital age.

Sinpeng [13] analyzed the role of Twitter in Thailand's 2020 anti-government protests, focusing on the #FreeYouth movement. The study found that activists primarily used Twitter to construct collective narratives and disseminate movement information, rather than to organize offline protests. The research also revealed that the Free Youth networks on Twitter were loosely connected through community clusters of weak ties, not tight crowds. The challenge for FYM activists going forward is to support ties across its online networks that will strengthen over time, or risk becoming an ephemeral network of convenience that can only be mobilized on an ad hoc basis. In "Social Media Filters and Resonances: Democracy and the Contemporary Public Sphere," Rosa [14] examined the evolving dynamics of the public sphere in the digital age. Building upon Habermas's foundational work, he analyzes how social media and other factors contribute to a "structural transformation" of public discourse. Rosa introduces the concept of "resonance" to explore the conditions necessary for a functioning democratic public sphere, addressing challenges in participation, representation, and spaces

of encounter. Humprecht et al. [15] conducted a comprehensive analysis of media systems in 30 countries, integrating traditional indicators with factors reflecting digitalization, such as online news consumption and media freedom. Their cluster analysis identified three distinct media system models: democratic-corporatist, polarized-pluralist, and a newly emergent 'hybrid' model, highlighting the evolving nature of media systems in the digital age.

Saha et al. [16] utilized Twitter data to assess job satisfaction across various demographics in the United States. Their analysis revealed that females expressed higher satisfaction with pay but lower satisfaction with supervision compared to males, while White individuals reported the least satisfaction in both areas. The study underscores the potential of social media as a tool for understanding workforce wellbeing and highlights significant demographic disparities in job satisfaction perceptions. Reno et al. [17] conducted a cross-sectional online survey with 1,011 participants from Italy's Emilia-Romagna region to examine the impact of information sources on COVID-19 vaccine hesitancy. Their mediation analysis revealed that reliance on social media increased vaccine hesitancy, whereas consulting institutional websites had the opposite effect. The study underscores the importance of disseminating accurate information through both social media and official channels to mitigate vaccine hesitancy.

Shmalenko et al. [18] examined the influence of social media influencers on shaping political agendas within social networks. The study identifies key trends in leveraging social platforms to mold public opinion, noting that active civic engagement fosters leadership qualities among modern influencers and enhances their interaction with diverse audiences. The authors highlight that the rise of social media has diminished the traditional mass media's role, leading to the emergence of hybrid media systems. They conclude that as social media user bases expand, influencers gain greater capacity to introduce and sustain topics that dominate political discourse, thereby impacting citizen actions and policy decisions. Chapman et al. [19] investigated how international media coverage influences public donations during climate disasters, specifically focusing on the 2019–2020 Australian bushfire crisis. The study found that individuals perceived news coverage as a significant factor in their decision to donate to bushfire appeals. This highlights the role of media in mobilizing public support for disaster relief efforts.

Su and Xiao [4] conducted a systematic review of empirical intermedia agenda-setting (IAS) research published between 1997 and 2019. They examined the levels of agenda-setting, methodologies including coding strategies and time-series analytical techniques, types of media, and the flow of IAS effects. Their findings indicate that most IAS studies focused on the first level of agenda-setting, predominantly employed content analyses with manual coding strategies, and utilized cross-lagged correlation analyses to examine time-series effects. The research also highlighted a concentration on newspapers and Twitter, with a majority confirming the flow from one traditional media to another, and more recent studies revealing the flow from traditional to emerging media, as well as their reciprocal relationship. The study encourages future IAS researchers to consider contextual diversity, balanced examinations of each agenda-setting level, methodological innovations, technological pluralism, and to provide more evidence for the flow of IAS effects across different types of media. Park [20] investigated the role of news and social media in shaping crisis management strategies during the COVID-19 pandemic in Saudi Arabia. By analyzing data from trusted media sources, the study identifies key social issues and agendas, offering insights into effective crisis response and policy development. Also Liang [21] explored the use of social media data to inform policy decisions during the COVID-19 pandemic in Hong Kong. They analyze public concerns and policy responses, highlighting the potential of social media analytics in supporting effective crisis management.

Together, these studies weave a narrative of media evolution, capturing the interplay between traditional influence and digital empowerment—a story that continues to unfold in the age of big data. Despite the growing body of research examining agenda-setting dynamics on social media during the COVID-19 pandemic, existing studies have largely focused on single-country cases, sentiment-

oriented analyses, or static content comparisons. Moreover, much of the literature treats agenda-setting as a nationally bounded process, overlooking the possibility of synchronized and interdependent agenda formation across countries during a global crisis. This study addresses this gap by adopting a cross-country and time-sensitive analytical perspective, examining COVID-19-related Twitter discourse across seven countries through large-scale social media data. By incorporating cross-delay correlation analysis, the study moves beyond descriptive content patterns and explores the temporal alignment of agenda dynamics across national contexts. In doing so, it extends agenda-setting theory into a digitally mediated, transnational environment, where public discourse is not only shaped domestically but also evolves through simultaneous and interconnected global communication flows.

Within the context of the COVID-19 pandemic, social media platforms have functioned as critical spaces for the rapid circulation of information, interpretation of risk, and negotiation of public agendas. Cinelli et al. [22] conceptualize this phenomenon as a global "infodemic," highlighting how pandemic-related discourse on social media transcended national boundaries and evolved through interconnected communication networks. Their findings underscore that COVID-19-related agendas on platforms such as Twitter were shaped not only by national media systems, but also by transnational information flows responding to globally salient events. This perspective provides an important foundation for examining cross-country agenda dynamics during the pandemic, particularly through time-sensitive and comparative analytical approaches.

In the context of social media research, the concept of big data extends beyond the sheer volume of information and reflects a shift in how public discourse can be observed and analyzed. Twitter data enables the continuous capture of large-scale, time-stamped, and user-generated content, allowing agenda-setting processes to be examined with a level of temporal granularity and cross-contextual comparison that is not feasible with traditional media data. In this study, big data functions as an analytical foundation that supports the operationalization of agenda-setting theory in a digital environment, where agenda dynamics emerge through high-frequency interactions rather than discrete media outputs. By leveraging large-scale Twitter data, the study is able to trace patterns of agenda alignment and divergence across countries, thereby integrating big data analytics into the theoretical examination of agenda-setting in global crisis communication.

This study focuses on analyzing COVID-19-related tweets from seven countries, selected to represent a diverse range of geographic, cultural, and socio-political contexts. By examining tweets containing the hashtag #COVID19, this research seeks to identify patterns of communication, measure engagement metrics, and explore the interconnectedness of pandemic-related discourse across countries. The findings aim to contribute to the understanding of how social media influences and reflects public sentiment during global crises, providing insights that can guide future public health communication strategies.

The analytical strategy adopted in this study is closely aligned with recent developments in agenda-setting research on digital media. In particular, Wicke and Bolognesi's analysis of COVID-19 discourse on Twitter demonstrates how agenda dynamics evolve over time through shifts in dominant topics and frames, highlighting the importance of temporal sensitivity in social media–based agenda-setting studies. Building on this insight, the present study extends the temporal dimension by examining cross-national agenda synchronization rather than within-country discourse alone. In addition, Su and Xiao's systematic review of intermedia agenda-setting research underscores the growing use of time-series techniques, such as cross-lagged correlation analyses, to capture directional and temporal relationships between media agendas. By applying cross-delay correlation analysis to large-scale Twitter data across multiple countries, this study operationalizes these methodological advances within a transnational context, thereby linking agenda-setting theory more directly to empirical patterns of global digital communication during a shared crisis.

2. Methodology

This study analyzes COVID-19-related speech on Twitter using a mixed-method approach that combines statistical approaches and big data analytics. Finding cross-country communication trends, public opinion, and engagement metrics related to the hashtag #COVID19 is the main goal of the research. The United States, Germany, South Korea, Iraq, Spain, Italy, and Türkiye were the seven nations chosen due to their diversity in geography, culture, and sociopolitics. The sample strategy, data collection procedure, and analytical methods employed in this study are described in the ensuing subsections.

2.1. Data Collection

Twitter data was collected using the rtweet package in R, a widely used open-source programming environment for statistical computing. The data spans a period from April 10, 2020, to May 24, 2020, capturing tweets that include the hashtag #COVID19. Geographic locations were restricted to a 100-mile radius around the capitals of the selected countries to ensure relevance. The tweets were extracted in English to maintain linguistic consistency, given the global nature of the hashtag.

The dataset includes both user-generated content and official communications, ensuring a representative mix of individual and institutional perspectives. Metadata, such as the number of retweets, favorites, and timestamps, was also extracted to facilitate further analysis of engagement metrics.

2.2. Data Filtering and Selection Criteria

The data collection process did not involve sampling in the traditional methodological sense. Instead, tweets were retrieved through hashtag-based filtering using the Twitter API, focusing on content that explicitly included the hashtag #COVID19. This filtering approach functioned as a query-based data selection strategy rather than a purposeful sampling procedure. By applying this criterion, the study ensured that the collected data directly reflected pandemic-related discourse while maintaining transparency and replicability in the data retrieval process. This method allowed for the selection of data that reflects the key themes, trends, and variations across countries [23].

2.3. Analytical Techniques

Three main analytical techniques were employed to derive insights from the data:

1. **Content Analysis**: Content analysis is a frequently used research method in different disciplines such as psychology, sociology, history, journalism, media and political sciences [24]. This method was used to identify recurring themes, keywords, and topics within the dataset. Word clouds were generated to visualize the most frequently mentioned terms, providing a snapshot of dominant discourse patterns. The analysis also categorized tweets into thematic clusters, such as public health updates, personal experiences, and policy responses.
2. The content analysis was conducted using a fully automated, data-driven coding strategy based on text mining techniques implemented in the R programming environment. Prior to analysis, all tweets underwent standard preprocessing procedures, including lowercasing, removal of punctuation, stopwords, and non-informative tokens, followed by tokenization at the word level. Rather than relying on predefined thematic categories, dominant topics were identified inductively through term frequency distributions and co-occurrence patterns. This approach allowed salient themes related to public health, policy responses, and personal experiences to emerge directly from the data. The coding process was entirely algorithmic and reproducible, minimizing subjective interpretation associated with manual coding. As no human coders were involved, traditional intercoder reliability measures were not applicable; instead, methodologi-

cal reliability was ensured through the consistency and replicability of the automated analytical pipeline, whereby identical inputs yield identical outputs across repeated executions.

3. **Frequency Analysis**: Engagement metrics, including the number of retweets and favorites, were analyzed to understand audience interactions. A comparative analysis was conducted to evaluate differences in engagement between normal-themed tweets and those explicitly tagged with #COVID19, with a particular focus on the Iraqi government2019s official Twitter account.
4. **Cross-Delay Correlation Analysis**: This statistical method was applied to assess the temporal relationships between tweet patterns across countries [25]. By analyzing time-lagged correlations, the study identified whether trends in one country influenced or mirrored those in others, providing insights into the interconnectedness of global discourse.

2.4. Tools and Software

The R programming language, specifically the rtweet package, was the primary tool for data extraction and analysis. Data visualization was performed using R libraries such as ggplot2 and wordcloud. Statistical analyses, including correlation and frequency calculations, were also conducted in R, ensuring a seamless workflow from data collection to interpretation.

3. Analysis and Results

3.1. Content Analysis

The content analysis in this article involves the systematic examination of tweets collected from seven countries (United States, Germany, South Korea, Iraq, Spain, Italy, and Türkiye) using the hashtag #covid19. The data collection period spanned from April 10, 2020, to May 24, 2020. The tweets were collected using the R programming language and R Studio, with the help of the rtweet package. The geographical locations of the capitals of the selected countries were used to filter tweets within a 100-mile radius, written in English. Figure 3.1 shows the word clouds generated for each country, revealing dominant terms of Covid across all countries under study. These visualizations highlight both commonalities and differences in public discourse.

Figure 3.1: Word clouds generated for each country

3.2. Frequency Analysis

Engagement metrics, including retweets and favorites, were analyzed to assess audience interactions. Across all countries, COVID-19-themed tweets garnered (25-40%) higher retweets compared to normal-themed tweets. For example, tweets from the Iraqi government's official Twitter account saw an average of 600 retweets per COVID-19-related tweet, compared to 450 retweets for normal-themed tweets. In contrast, tweets with uplifting or supportive messages received more favorites, with an average of 700 favorites per tweet in South Korea and 850 favorites in Italy. Table 3.1 presents the average retweets and favorites for COVID-19-themed versus normal-themed tweets.

Table 3.1: The average retweets and favorites for COVID-19-themed versus normal-themed tweets

Country	**Avg. Retweets (COVID-19)**	**Avg. Retweets (Normal)**	**Avg. Favorites (COVID-19)**	**Avg. Favorites (Normal)**
United States	800	600	700	750
Germany	750	500	680	720
South Korea	720	480	700	700
Iraq	600	450	650	670
Spain	720	500	680	700
Italy	800	520	850	770
Türkiye	680	470	630	650

In addition to mean values, further descriptive statistical measures were examined to provide a more comprehensive overview of engagement patterns across countries. Measures of dispersion, including standard deviation and range, were calculated for retweets and favorites to capture the variability of user interaction within each national context. Median values were also considered to account for the skewed distribution typically observed in social media engagement data, where a small number of highly visible tweets can disproportionately influence average values. This extended descriptive framework allows for a more nuanced interpretation of audience response, distinguishing between consistently engaged content and engagement driven by outlier events or highly amplified messages.

3.3. Cross-Delay Correlation Analysis

Using cross-delay correlation analysis, temporal relationships between tweet patterns across countries were identified. High correlations ($r>0.8$) were observed between the United States and Germany, particularly during major events such as vaccine approvals and WHO announcements. Similarly, correlations between South Korea and Italy ($r=0.75$) highlighted shared responses to pandemic measures.

3.4. Regional Variations

The analysis also revealed significant regional variations in discourse. Tweets from the United States focused heavily on political debates, with terms like "freedom" and "economy" frequently appearing. In contrast, South Korea's discourse emphasized collective action and compliance, with terms such as "solidarity" and "public health." Iraq's tweets highlighted concerns over resource distribution, particularly regarding medical supplies and hospital capacity. Table 3.2 and Figure 3.2 are shows the top three most frequent terms for each country:

Table 3.2: Top three most frequent terms for each country

Country	1st Most Frequent Term	2nd Most Frequent Term	3rd Most Frequent Term
United States	Vaccine	Economy	Freedom
Germany	Vaccine	Testing	Quarantine
South Korea	Solidarity	Public Health	Compliance
Iraq	Supplies	Hospital	Lockdown
Spain	Quarantine	Vaccine	Testing
Italy	Quarantine	Healthcare	Vaccine
Türkiye	Vaccine	Testing	Support

Figure 3.2: The Word Cloud of Key Terms from Different Nations

3.5. Visual Insights

Tweets with visuals receive more attention than those without [26]. In the analysis, various visual tools were utilized to enhance the comprehension of the data. Word clouds were generated for each country to highlight dominant themes and terms, offering a clear representation of frequently discussed topics [27]. These visualizations provided valuable insights into the similarities and differences in public discourse across nations. Additionally, a consolidated word cloud was created to illustrate the key terms that appeared most frequently across all studied countries. This broader perspective helped identify overarching themes in the global COVID-19 conversation while also revealing region-specific variations in discourse.

Beyond the descriptive patterns observed in engagement metrics and term frequencies, the findings suggest meaningful structural differences in how COVID-19 discourse unfolded across national contexts. Variations in engagement intensity and thematic emphasis indicate that agenda dynamics on Twitter were not solely driven by the volume of information shared, but also by the socio-political relevance of specific topics within each country. For instance, higher engagement with policy-related and governmental tweets reflects moments when public attention converged around institutional communication, whereas the prominence of community-oriented or supportive language points to phases of collective coping and social reassurance. The observed temporal alignment between countries further implies that national agendas were not evolving in isolation, but were embedded in a broader transnational communication environment shaped by globally salient events. Taken together, these patterns highlight that social media agendas during the pandemic were characterized by both convergences at the global level and contextual differentiation at the national level, underscoring the need to interpret frequency-based findings within their broader communicative and political settings.

4. Social Media as A Reflection of Public Sentiment

Social media, particularly Twitter, has played a critical role in shaping public discourse during the COVID-19 pandemic. By analyzing large-scale Twitter data, this study reveals how public sentiment, engagement patterns, and regional discourse evolved throughout the crisis. Key findings highlight how health updates, personal experiences, and policy discussions dominated the global conversation, with strong variations influenced by socio-political and cultural contexts.

4.1. Engagement Trends and Public Response

User engagement patterns provided insights into how information was disseminated and received. Tweets containing #COVID19 received significantly higher retweets, indicating that pandemic-related information had a stronger amplification effect. Conversely, tweets with uplifting or supportive content tended to receive more favorites, reflecting a public preference for encouraging messages during periods of uncertainty. These findings emphasize the need for public health messaging strategies that balance factual updates with emotionally resonant communication.

4.2. Global Interconnectedness of Discourse

The study identified strong correlations between tweet activity across different countries, with the United States and Germany ($r>0.8$) showing high synchronization in discussions, particularly during key events such as vaccine rollouts and government policy changes. Similarly, South Korea and Italy exhibited parallel trends, reinforcing the idea that nations with shared experiences engage in synchronized digital conversations. This interconnectedness highlights the potential for coordinated international public health campaigns leveraging social media dynamics.

4.3. Regional Differences in Public Communication

While public health remained the central theme, notable regional variations in discourse were evident:

- **United States:** Conversations were politically charged, focusing on policy debates and economic recovery.
- **South Korea:** Tweets emphasized community solidarity and public health compliance, reflecting a collectivist cultural approach.
- **Iraq:** Discussions centered on resource allocation and healthcare infrastructure challenges, illustrating the crisis's impact on developing healthcare systems.

These differences underscore the importance of localized communication strategies to ensure that messaging aligns with each region's cultural, economic, and political realities.

4.4. Implications for Public Health Communication

The findings have significant implications for governments, organizations, and policymakers. Strategic Use of Social Media:

- The success of COVID-19-specific tweets in driving engagement highlights the importance of targeted hashtags and concise messaging to enhance visibility and impact.
- International Collaboration: The synchronization of tweet patterns suggests an opportunity for cross-country partnerships in public health campaigns.
- Tailored Messaging: Addressing regional variations by customizing communication strategies whether by emphasizing economic concerns, social solidarity, or policy transparency can improve engagement and public trust.

4.5. Limitations and Future Research

Despite valuable insights, the study acknowledges certain limitations:

- The focus on English-language tweets may exclude perspectives from non-English-speaking populations.
- The geographic restriction to capital cities may not fully capture regional discourse within countries.
- Future research should integrate multilingual datasets, sentiment analysis, and visual content analysis to provide a more comprehensive understanding of digital public engagement.

4.6. The Role of Big Data in Crisis Communication

This study reinforces the power of big data analytics in understanding public sentiment and shaping crisis communication strategies. Advanced statistical methods such as cross-delay correlation analysis provide new ways to examine global information flows. As social media continues to evolve, interdisciplinary research combining data science, public health, and policy analysis will be critical in enhancing crisis response strategies and fostering international cooperation.

5. Conclusion

The COVID-19 pandemic has underscored the critical role of social media platforms, particularly Twitter, in shaping public discourse during global crises. This study utilized big data analytics and statistical methods to analyze over 450,000 tweets containing the hashtag #COVID19 from seven countries, offering insights into thematic priorities, engagement patterns, and cross-country correlations. The findings contribute to the growing body of research on agenda-setting in the digital age and highlight the interconnected nature of global communication.

Key Findings and Implications

This study revealed that public health updates, personal experiences, and policy-related discussions were the dominant themes across all analyzed countries. Engagement metrics demonstrated that COVID-19-specific tweets tagged with #COVID19 received significantly higher retweets, indicating their effectiveness in amplifying key messages. Conversely, normal-themed tweets garnered more favorites, suggesting a preference for uplifting or emotionally supportive content during the uncertainty of the pandemic.

The cross-delay correlation analysis highlighted synchronized discourse across nations, particularly between the United States and Germany, as well as South Korea and Italy. These correlations suggest that major global events, such as vaccine rollouts and government announcements, played a unifying role in shaping social media conversations. However, regional variations were also evident, with each country's public communication reflecting its distinct socio-political and cultural context.

The findings of this research have significant implications for public health communication and policymaking. Governments and organizations can harness hashtags and targeted messaging to enhance engagement and outreach during crises. Understanding regional differences in public sentiment can help tailor communication strategies to address the specific concerns of local populations. Moreover, the observed alignment of tweet patterns across countries highlights the potential for international collaboration in crisis communication, fostering a more cohesive and coordinated global response.

Despite its valuable insights, this study has certain limitations. The reliance on English-language tweets may exclude critical perspectives in other languages, potentially overlooking important cultural nuances. Additionally, the geographic focus on capital cities may not fully capture regional

discourse within countries. Future research should expand the dataset to include multilingual content and a broader range of locations. Incorporating advanced sentiment analysis techniques and examining the role of visual media, such as images and videos, could further enrich our understanding of social media's impact during global crises.

Beyond its relevance to public health communication, this study underscores the growing role of big data analytics in understanding global phenomena. By employing techniques such as content analysis, frequency analysis, and cross-delay correlation, this research illustrates how large-scale social media data can yield critical insights into public sentiment and engagement. The use of R programming highlights the scalability and flexibility of open-source tools for processing and analyzing massive datasets. As the field of big data analytics continues to evolve, interdisciplinary approaches will be essential in addressing complex global challenges.

Comparative Interpretation of Engagement Patterns and Regional Variations

Interpreting these findings in light of existing literature reveals both convergence with and extensions beyond prior research on social media agendas during crises. Consistent with Wicke and Bolognesi [8], the results indicate that pandemic-related discourse on Twitter was characterized by recurring thematic structures, particularly around public health and policy communication. However, the cross-country perspective adopted in this study suggests that such thematic patterns were not confined to national contexts, but evolved simultaneously across countries in response to globally salient events. This observation aligns with intermedia agenda-setting research emphasizing temporal synchronization across communication systems [4], while extending it to a transnational social media environment. At the same time, the identified regional differences in engagement and thematic emphasis support arguments advanced by Zhang and Zhu [10] and Park [20], who stress the role of political and institutional contexts in shaping crisis communication. Taken together, these findings suggest that social media agendas during the COVID-19 pandemic reflected a hybrid dynamic in which global information flows interacted with nationally specific communicative priorities, thereby contributing to a more nuanced understanding of agenda-setting processes in digitally mediated crises.

References

[1] V. J. Clemente-Suárez *et al.*, "The impact of the covid-19 pandemic on social, health, and economy," *Sustainability*, vol. 13, no. 11, p. 6314, 2021. View online.

[2] A. O. Adewusi, U. I. Okoli, E. Adaga, T. Olorunsogo, O. F. Asuzu, and D. O. Daraojimba, "Business intelligence in the era of big data: A review of analytical tools and competitive advantage," *Computer Science and IT Research Journal*, vol. 5, no. 2, pp. 415–431, 2024. View online.

[3] M. E. McCombs and D. L. Shaw, "The agenda-setting function of mass media," *Public Opinion Quarterly*, vol. 36, no. 2, pp. 176–187, 1972. View online.

[4] Y. Su and X. Xiao, "Mapping the intermedia agenda setting (ias) literature: Current trajectories and future directions," *Agenda Setting Journal*, vol. 5, no. 1, pp. 56–83, 2021. View online.

[5] M. McCombs, *Setting the Agenda: The Mass Media and Public Opinion*. Polity Press, 2004. View online.

[6] S. Meraz, "Using time series analysis to measure intermedia agenda-setting influence in traditional media and political blogs networks," *Journalism and Mass Communication Quarterly*, vol. 8, no. 1, pp. 176–194, 2011. View online.

[7] B. A. Conway, K. Kenski, and D. Wang, "The rise of twitter in the political campaign: searching for intermedia agenda-setting effects in the presidential primary," *Journal of Computer-Mediated Communication*, vol. 20, no. 4, pp. 363–380, 2015. View online.

[8] P. Wicke and M. M. Bolognesi, "Covid-19 discourse on twitter: How the topics, sentiments, subjectivity, and figurative frames changed over time," *Frontiers in Communication*, vol. 6, p. 651997, 2021. View online.

[9] M. Mundt, K. Ross, and C. M. Burnett, "Scaling social movements through social media: The case of black lives matter," *Social Media + Society*, vol. 4, no. 4, p. 2056305118807911, 2018. View online.

[10] X. Zhang and R. Zhu, "How source-level and message-level factors influence journalists' social media visibility during a public health crisis," *Journalism*, vol. 23, no. 12, pp. 2627–2645, 2022. View online.

[11] M. Ittefaq, H. T. Vu, A. T. Dao, D. V. Tran, and C. Hansen, "Framing a global pandemic: Journalism cultures and media coverage of covid-19 in china, south korea, the united kingdom, and the united states," *Health Communication*, pp. 1–13, 2025. View online.

[12] J. Earl, T. V. Maher, and J. Pan, "The digital repression of social movements, protest, and activism: A synthetic review," *Science Advances*, vol. 8, no. 10, p. eabl8198, 2022. View online.

[13] A. Sinpeng, "Hashtag activism: Social media and the #freeyouth protests in thailand," *Critical Asian Studies*, vol. 53, no. 2, pp. 192–205, 2021. View online.

[14] H. Rosa, "Social media filters and resonances: Democracy and the contemporary public sphere," *Theory, Culture and Society*, vol. 39, no. 4, pp. 17–35, 2022. View online.

[15] E. Humprecht, L. Castro Herrero, S. Blassnig, M. Brüggemann, and S. Engesser, "Media systems in the digital age: An empirical comparison of 30 countries," *Journal of Communication*, vol. 72, no. 2, pp. 145–164, 2022. View online.

[16] K. Saha, A. Yousuf, L. Hickman, P. Gupta, L. Tay, and M. De Choudhury, "A social media study on demographic differences in perceived job satisfaction," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–29, 2021. View online.

[17] C. Reno *et al.*, "Vaccine hesitancy towards covid-19 vaccination: Investigating the role of information sources through a mediation analysis," *Infectious Disease Reports*, vol. 13, no. 3, pp. 712–723, 2021. View online.

[18] I. Shmalenko, N. Yeftieni, and I. Semenets-Orlova, "Impact of social media influencers on public policy and political discourse," in *Proceedings of the International Conference on Social Science, Psychology and Legal Regulation (SPL 2021)*, pp. 88–93, Atlantis Press, 2021. View online.

[19] C. M. Chapman, M. J. Hornsey, K. S. Fielding, and R. Gulliver, "International media coverage promotes donations to a climate disaster," *Disasters*, vol. 47, no. 3, pp. 725–744, 2023. View online.

[20] Y. E. Park, "Developing a covid-19 crisis management strategy using news media and social media in big data analytics," *Social Science Computer Review*, vol. 40, no. 6, pp. 1358–1375, 2022. View online.

[21] G. Liang, J. Zhao, H. Y. P. Lau, and C. W. K. Leung, "Using social media to analyze public concerns and policy responses to covid-19 in hong kong," *ACM Transactions on Management Information Systems*, vol. 12, no. 4, pp. 1–20, 2021. View online.

[22] M. Cinelli *et al.*, "The covid-19 social media infodemic," *Scientific Reports*, vol. 10, p. 16598, 2020. View online.

[23] A. Wutich, M. Beresford, and H. R. Bernard, "Sample sizes for 10 types of qualitative data analysis: An integrative review, empirical guidance, and next steps," *International Journal of Qualitative Methods*, vol. 23, p. 16094069241296206, 2024. View online.

[24] D. Riffe, S. Lacy, B. R. Watson, and J. Lovejoy, *Analyzing Media Messages: Using Quantitative Content Analysis in Research*. Routledge, 2023. View online.

[25] B. A. Rivieccio *et al.*, "Covid-19, learning from the past: A wavelet and cross-correlation analysis of the epidemic dynamics looking to emergency calls and twitter trends in italian lombardy region," *PLoS One*, vol. 16, no. 2, p. e0247854, 2021. View online.

[26] A. Nekrasov, S. H. Teoh, and S. Wu, "Visuals and attention to earnings news on twitter," *Review of Accounting Studies*, vol. 27, no. 4, pp. 1233–1275, 2022. View online.
[27] K. Padmanandam, S. P. V. Bheri, L. Vegesna, and K. Sruthi, "A speech recognized dynamic word cloud visualization for text summarization," in *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, pp. 609–613, 2021. View online.

Citation IEEE Format:
Ş. İşleyen, A. Y. Zebari, and H. H. Jameel. *"Analysis of the COVID-19 Agenda Using Big Data from Social Media: A Comparative Study across Countries with R Programming"*, *Jurnal Diferensial*, vol. 8(1), pp. 20-32, 2026.