Article

A Retrospective Analysis of the COVID-19 Infodemic in Saudi Arabia

Ashwag Alasmari [1,*,†], Aseel Addawood [2,†], Mariam Nouh [3], Wajanat Rayes [4] and Areej Al-Wabil [5]

1 Computer Science Department, King Khalid University, Abha 62529, Saudi Arabia
2 Information System Department, Imam Mohammad Bin Saud University, Riyadh 11564, Saudi Arabia; aasdawood@imamu.edu.sa
3 Center for Complex Engineering Systems (CCES) at KACST and MIT, King Abdulaziz City for Science and Technology, Riyadh 12354, Saudi Arabia; mnouh@kacst.edu.sa
4 Department of Information Science, Umm Al-Qura University, Makkah 21955, Saudi Arabia; warayes@uqu.edu.sa
5 College of Engineering, Alfaisal University, Riyadh 11533, Saudi Arabia; awabil@alfaisal.edu
* Correspondence: aasmry@kku.edu.sa
† These authors contributed equally to this work.

Abstract: COVID-19 has had broad disruptive effects on economies, healthcare systems, governments, societies, and individuals. Uncertainty concerning the scale of this crisis has given rise to countless rumors, hoaxes, and misinformation. Much of this type of conversation and misinformation about the pandemic now occurs online and in particular on social media platforms like Twitter. This study analysis incorporated a data-driven approach to map the contours of misinformation and contextualize the COVID-19 pandemic with regards to socio-religious-political information. This work consists of a combined system bridging quantitative and qualitative methodologies to assess how information-exchanging behaviors can be used to minimize the effects of emergent misinformation. The study revealed that the social media platforms detected the most significant source of rumors in transmitting information rapidly in the community. It showed that WhatsApp users made up about 46% of the source of rumors in online platforms, while, through Twitter, it demonstrated a declining trend of rumors by 41%. Moreover, the results indicate the second-most common type of misinformation was provided by pharmaceutical companies; however, a prevalent type of misinformation spreading in the world during this pandemic has to do with the biological war. In this combined retrospective analysis of the study, social media with varying approaches in public discourse contributes to efficient public health responses.

Keywords: infodemic; COVID-19 misinformation; social media; policy intervention; Saudi Arabia

1. Introduction

Social media screening for public health has emerged as an essential component in combating the COVID-19 pandemic. False rumors, misinformation, and disinformation (e.g., "fake news") are diffused and often inadvertently endorsed through social media platforms markedly faster, deeper, and more broadly than trustworthy information. Research has shown that misinformation can foster an atmosphere of panic and discrimination in pandemics [1]. Hence, a need for the public during pandemics to have access to clear, up-to-date information and to be transparent in decision-making at the strategic and operational levels. With the global growth of social media platforms, there are questions regarding how regional cultural factors shape online engagement, especially in the context of infodemics [2], a term used to describe the overabundance of information, both online and offline. In February 2021, the World Health Organization announced that the coronavirus pandemic accompanied an 'infodemic' of misinformation (WHO 2020). In this study, we examine the COVID-19 infodemic in the context of one country, Saudi Arabia.

Saudi Arabia has one of the most significant social media presences in the world, possibly due to the country's relatively high rate of smartphone ownership when compared to other countries in the world. In Saudi Arabia, there are about 40.20 million mobile subscribers, resulting in a mobile penetration rate of 116 percent of the total population [3]. Saudi Arabia has quickly evolved from a media landscape characterized by print and television to one dominated by online engagement over the last two decades. Twitter, in particular, has been considered a valuable source of news in both Arabic and English and a public medium for expression for residents of Saudi Arabia during the COVID-19 pandemic [4].

Misinformation about Arabic content on social media was causing mistrust between the general public and health officials. It promoted stigma that could consequently prevent affected persons with COVID-19 from seeking medical attention, thus perpetuating disease transmission and creating friction in the local response [5–8]. This is evidenced by recent studies that emerged from the local circumstances (e.g., [7,9]).

This study aims to gain insights into information behavior when discussing the pandemic on a widely used social media channel, Twitter. We are specifically interested in identifying misinformation spread regarding epidemics particularly in Saudi Arabia. Questions to be answered include:

- What forms of misinformation spread the most in pandemics?
- How does misinformation evolve over time?

This work is part of an ongoing effort to capture and classify misinformation, which has emerged as an essential area of technical and social research in the context of infodemics. In addition, the study findings can help shape effective public health communication to support efforts to reduce the effects of misinformation.

2. Background and Related Work

Systematic studies of misinformation on social media platforms and digital social listening for public health predate the COVID-19 era, tracing back to outbreaks such as Ebola [10] and H1N1 [11]. The spread of misinformation on social media has shown to have multiple negative wide-scale effects. It can manipulate public opinion [12] and incite fear and chaos [13] that lead to several social disorders. Such phenomena have a more prominent negative effect during a global health crisis and pandemic. As a result, a timely public health response to any emerging concern during pandemics can limit the spread of misinformation and prevent public panic; social media present a rich source of information that should be harnessed to support the public health response during pandemics, as they provide real-time access to community beliefs [14].

Many countries, including Saudi Arabia, have issued new regulations and laws to manage the spread of misinformation. The Ministry of Interior Affairs announced on 5 May 2020 that anyone who disseminated any misinformation regarding COVID-19 on social media that could cause panic in any form or lead to a violation of precautionary measures would be liable to fines or a maximum prison sentence of five years [15].

2.1. Health Misinformation in the Arabic Language

During the pandemic the use of Twitter has increased as it provided a medium to discuss topics related to COVID-19, including misinformation topics. There are several spoken languages which are heavily used on social media platforms. However, not all these languages have easy access to verified or reliable information, especially during this tough time. There are many studies investigating misinformation on Twitter related to COVID-19 in different languages (e.g., [1,9,16–18]). For example, Jussila et al. [17] explored misinformation related to COVID-19 in Finnish.

Apart from the difficulties presented by the spread of misinformation, the Arabic language poses challenges primarily due to the lexical variation of different Arabic dialects. As a result, that makes misinformation exist in more than one dialect, making it difficult to detect. Hence, developing systems capable of automatically detecting misinformation

in Arabic content is urgent. Some of these efforts concerning the Arabic language were made by Haouari et al. [16], who collected an Arabic dataset from Twitter that supported verification for both the claim and tweet level. Another study by Alqurashi et al. [18] experimented with varying machine learning classifiers and word embeddings to auto classify misinformation in Arabic. Despite the insights provided by these studies, there is still a need for more studies which focus on topics and insights concerning misinformation spread in Arabic on Twitter with respect to Saudi Arabia. This paper aimed to investigate the type of misinformation spread in Saudi Arabia during the early phases of the COVID-19 outbreak. The narrative threads of misinformation that have circulated in Arabic-speaking populations since the COVID-19 crisis broke out vary from full-throated conspiracy 76 theories to unscientific health advice.

2.2. Arabic COVID-19 Twitter Datasets

It is well established that social media data can help gauge public opinion for a more human-centered design of communication and engagement strategies to address public concerns before they become widespread. As noted by the World Health Organization (WHO) in the Infodemic Management virtual conference in 2020, "This can help to reach citizens who are undecided or confused about adopting COVID-19 public health, and social measures, including vaccination [19]." In the first few months of the pandemic, several annotated Twitter datasets emerged in the public domain, as described in Table 1. Other crisis-related Twitter datasets in the Arabic language, such as Adel and Wang corpus [20], also provide insights into the technical approaches that have been incorporated in analyzing Arabic Twitter content.

Table 1. Public Annotated COVID-19 Datasets in Arabic language.

Dataset	Timeframe	Tweets
ArCov-19 [21]	January–June 2020	785,000
ArCorona [22]	21 February–31 March 2020	1,000,000
Addawood's Dataset [23]	January–April 2020	3,800,000
Alam [24]	January–April 2020	218
Alsudais Dataset [9]	December 2019–April 2020	1,048,575

2.3. Tools to Understand, Measure, and Control the COVID-19 Infodemic

While an infodemic cannot be eliminated, it can be managed with the right precautions. The World Health Organization has led the efforts in this field to guide relevant research and effective practices across the globe and define public health research needs in order to advance this field during the COVID-19 pandemic. The first WHO Infodemiology Conference held on 29 June 2020, and several conferences followed, which had a threefold contribution: (1) Understanding the multidisciplinary nature of infodemic management; (2) identifying current examples and tools to understand, measure, and control infodemics; and (3) building a public health research agenda to direct focus and investment in this emerging scientific field with the overall aim of establishing a community of practice and research. A framework for managing infodemics was proposed, which requires a transdisciplinary approach to address the problem's complexity, including several disciplines such as mathematics, digital health, data science, and social and behavioral sciences [25].

Similarly, UNESCO published policy briefs focusing on the increasing threat related to the spread of COVID-19 misinformation. The briefs analyze the types of misinformation, investigate how individuals, governments, and media platforms respond to this phenomenon, review actions to combat it, and assess risks associated with applied measures to limit its spread. Finally, they provide recommendations on how to respond to the crisis, taking into consideration human rights concerns such as freedom of expression and privacy [26]. One method used by the agency is to promote facts about the COVID-19 disease from credible sources and motivate people to be more critical towards the information they see

online using hashtags campaigns such as ThinkBeforeClicking, ThinkBeforeSharing, and ShareKnowledge [27,28].

Furthermore, numerous recent studies examined the impact of Twitter and other social media platforms on COVID-19 and infodemics. Chen et al. [29] studied the information and misinformation landscape over a year-long period of Twitter to characterize the spread on social media. They used clustering and topic modeling techniques to identify the major narratives, including health misinformation and conspiracies. They found that the echo chamber effect contributes to misinformation spread as users who share questionable content are clustered more closely in the network than others. Vargas et al. [30] explored the use of network analysis techniques to detect disinformation campaigns and generate features for distinguishing them from legitimate activities. They trained a binary classifier based on statistical features extracted from both networks. The results showed that coordination patterns could be helpful for providing evidence of disinformation activity.

3. Materials and Methods

In this paper, we utilized a data-driven approach. We set out to map the contours of misinformation in Saudi Arabia and contextualize it within the socio-religiopolitical information environment. Our research is grounded on a mixed-methods approach bridging quantitative and qualitative methods to determine if information-exchanging behaviors can be used to minimize the effects of emergent misinformation [31].

3.1. Data Collection

3.1.1. Twitter Data

Data was collected on Twitter from the beginning in December 2019 to 10 April 2020, using several keywords related to the pandemic in Arabic [23]. The dataset was collected by identifying a list of popular hashtags and keywords used mainly by the public in the local context of Saudi Arabia. The trending hashtags were identified as trending where they specifically discuss precautionary measures governments have applied. These include discussions of curfew, business closures, and travel restrictions. Appendix A lists popular hashtags used by the early public pandemic in Saudi Arabia along with English translations. Data was collected using Crimson Hexagon (https://www.crimsonhexagon.com/, accessed on 27 September 2021), which is a social media analytic platform that provides paid data stream access. This tool allowed the collection of 3.8 million tweets and retweets discussing the pandemic in Arabic.

3.1.2. Survey Data

In addition to Twitter data, we also created a short survey (https://forms.gle/7Qjuxc14KP9pYJJbA, accessed on 27 September 2021) to collect rumors and fake news that spread in the community. The survey asked the community to share any information they encountered that they suspected to be misinformation. The survey asked three questions, concerning the nature of the misinformation, the source of the misinformation (e.g., Twitter, WhatsApp, Instagram, etc.), and a link to the misinformation if applicable. We used a snowball sampling approach [32] to distribute the survey within local communities. The online survey was distributed using social media such as Twitter and different local WhatsApp groups.

3.2. Identifying Misinformation Themes and Keywords

We build on previous work on thematic categories of study in the context of infodemics in general, and COVID-19 in particular [33]. Our goal is to provide transparency about our process so that the strengths and weaknesses of different approaches can be straightforward.

3.2.1. Misinformation Themes

At first, a list of misinformation was identified by an iterative cycle of reviewing coverage of public announcements made by local authoritative channels, including the ministry of health and other Saudi official governmental websites, to fact check the COVID-19-related misinformation. We also looked at this website (http://norumors.net, accessed on 27 September 2021), an unofficial but popular Saudi website, to check all types of fact misinformation. A total of 30 misinformation items were identified, broadly categorized under seven themes. These misinformation themes were chosen to cover a wide range of domains. These themes are related to pharmaceutical companies, health advice, conspiracy theories, biological war, Arab immunity, perception of Islamophobia and the 5G network.

3.2.2. Data Segmentation

For each misinformation theme, an Arabic keyword list was developed that covers the meaning of that misinformation. The keyword lists were generated using synonyms and acronyms. For example, concerning the misinformation about biological wars, keywords like biological warfare and biological weapon were included in the list. We listed an average of eight keywords, which can cover the various wording variants for each misinformation item. Table 2 shows a list of keywords used to retrieve relevant tweets from the Twitter dataset that match the defined misinformation themes.

Table 2. Keywords used to retrieve relevant tweets.

Theme	Keywords
Pharmaceutical Companies	شركات الادوية
Health Advice	الثوم، بخار الماء، إستنشاق، الغرغرة ، عسل الامريكية الماء والملح ، حبة بركة ، ليمون، كركم
Conspiracy Theories	مؤامرة، تآمر، أؤمن بنظرية ، الصهيونية ، المافيا الامريكية، الماسونية
Biological War	سلاح بيولوجي، قنبلة بيولوجية ، وزارة الدفاع الامريكية، هندسة جينية ، مخطط، غرض عسكري حرب بيولوجيه، حرب عالية
Arab Immunity	مناعة ضد كورونا، مناعة العرب، العرب
Perception of Islamophobia	ضد الاسلام، القضاء على الإسلام، القضاء على المسلمين، اضطهاد، قمع المسلمين، تصفية المسلمين، الغضب الإلهي
5G Network	الجيل الخامس ، أشعة الجيل الخامس، قاتل صامت

A Python script was developed to segment the data to categorize each tweet in the Twitter dataset as discussing one of these misinformation items. The script counted the number of times the keywords appeared in the tweet for each misinformation item. If a tweet contained one or more of the misinformation keywords, it was classified as discussing that misinformation item. Table 3 shows the number of tweets for each misinformation item.

Table 3. Misinformation Themes.

Theme	# Tweets	# Annotated Tweets
Pharmaceutical Companies	101	101
Health Advice	14,320	1010
Conspiracy Theories	3060	255
Biological War	4467	898
Arab Immunity	163	163
Perception of Islamophobia	482	198
5G Network	92	92

3.2.3. Misinformation Labeling and Validation

The main goal of this research is to understand the distribution of misinformation in social media. To do that, we developed a codebook for annotating the tweets in each misinformation category. Each tweet was annotated either as misinformation (any tweet that confirms and believes the misinformation); no misinformation (any tweet that provides general factual information about the virus, questions, news, or denies the misinformation); or not related (ads or anything that appears to be incorrectly classified). For annotating the

collected dataset, we utilized the shared information on the official websites and the official Twitter accounts of the Ministry of Health and WHO as a source of credible information. The COVID-19 pre-checked facts have been obtained from different fact-checking websites to build a ground-truth database.

A subset of each category was chosen randomly; two annotators then went through the data to label each tweet. The final corpus consists of 2717 tweets. Table 3 shows misinformation themes, a number of tweets for each theme, and the number of annotated tweets. To validate the annotation of these tweets, both annotators agreed on 97% of the annotation.

4. Results

The results of our analysis of 2717 tweets for the five months are described in this section.

4.1. Misinformation in Social Media

First, we need to understand how much misinformation is represented in our dataset, which might give a holistic overview of how misinformation manifests in our daily social media interactions. Figure 1 shows that tweets that include health advice such as eating garlic or using lemon and hot water for minimizing the chances of getting COVID-19 were the most common form of misinformation. The belief that pharmaceutical companies are benefiting from the pandemic was the second-most common type of misinformation. Furthermore, one of the most frequent types of misinformation is a similar hypothesis for the origin of COVID-19, which includes the biological war against the world.
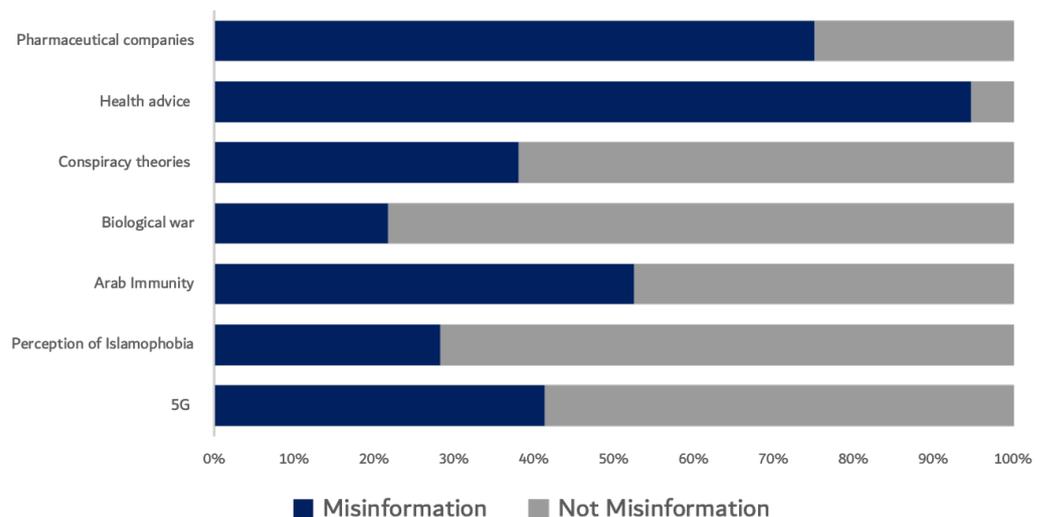

Figure 1. Distribution Of Misinformation Across Topics.

4.2. Types of Misinformation Emerging from Digital Social Listening during the COVID-19 Pandemic

In this study, we used word clouds to visualize the text corpus for each misinformation type to gain insights into the most frequent unigrams and bigrams. Figure 2 shows word clouds of the most frequent words associated with the seven types of misinformation. In general, the diagrams show that the most frequently occurring terms are نظريات المؤامرة "conspiracy theories", حرب بيولوجية "biological warfare", الشركات الطبية "pharmaceutical companies", الصين "China", السعودية "Saudi Arabia", مناعة "immunity". For many types of misinformation, terms related to anger and prayers show a high rate of occurrence.

(a) 5G Networks.

(b) Arab Immunity.

(c) Biological War.

(d) Conspiracy Theory.

(e) Health Advices.

(f) Perception of Islamophobia.

(f) Pharmaceutical Companies.

Figure 2. Word clouds for different types of misinformation.

4.3. Temporal Patterns in COVID-19 Related Digital Misinformation in Saudi Arabia

After annotating each tweet for thematic categories, as shown in Figure 3, we found that there was a rise in the amount of misinformation, especially following the second week of March. The momentum of misinformation had already started an upward trajectory before the Ministry of Health announced the lockdown in March 2020; however, it has continued to increase since that week.

Moreover, we can see that some misinformation items did not start circulating until later, while others have been spreading since the pandemic, especially misinformation related to pharmaceutical companies. The theme of health advice had the highest volume of all topics.
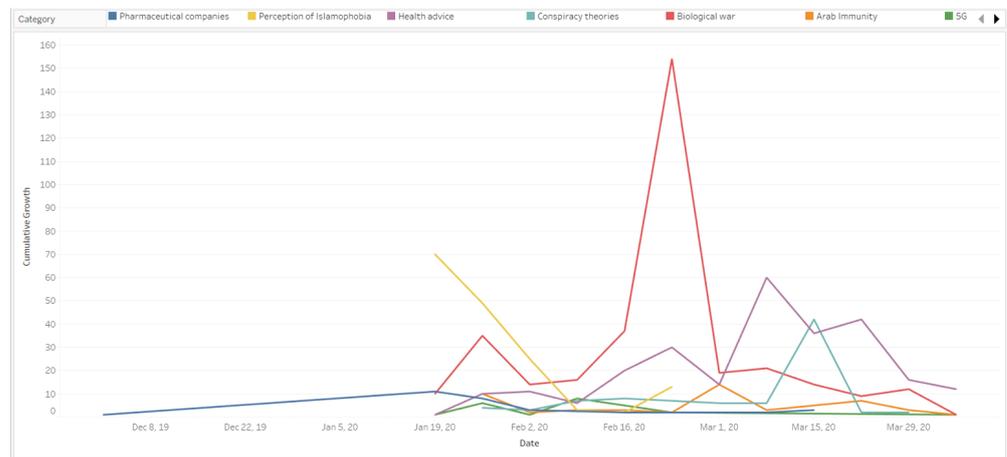
Figure 3. Cumulative Weekly Growth of Misinformation.

4.4. Community-Reported Misinformation-Survey

The survey was designed to measure the community and understand the topics and sources of the most common COVID-related misinformation experienced during the early stages of the pandemic.

A total of 88 respondents participated in the survey, which was available online and distributed through different social media sources. The majority of the survey respondents were between 21 and 40 years of age, and there were twice as many female respondents as male.

There were three sources of rumors reported by participants (See Figure 4): (1) Social circle (5%), i.e., through word of mouth from friends and family; (2) traditional media (8%) such as TV and newspapers; and (3) social media platforms (87%) such as Twitter and Facebook. Social media platforms were reported as the most common source of rumors as these platforms are nowadays the go-to media for information in general. This finding supports our work on focusing on social media platforms to understand the different types of misinformation emerging during the COVID-19 pandemic.
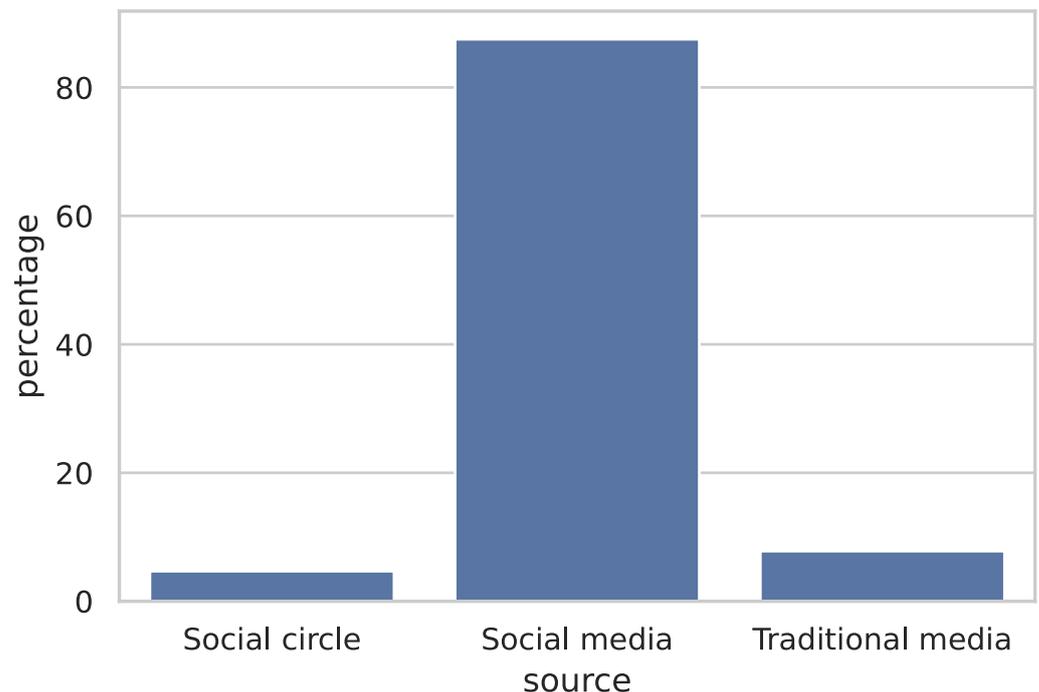

Figure 4. Reported Sources of Misinformation by the Community.

When examining the social media platform sources, we found that the most reported source of rumors in the community was WhatsApp (46%), followed by Twitter (41%). Other social media platforms such as Facebook, Instagram, YouTube, and Snapchat were also reported with lower frequencies as shown in Figure 5. This result is in line with the reported trend in Saudi Arabia of WhatsApp and Twitter being the most utilized and penetrated platforms in the country [3].
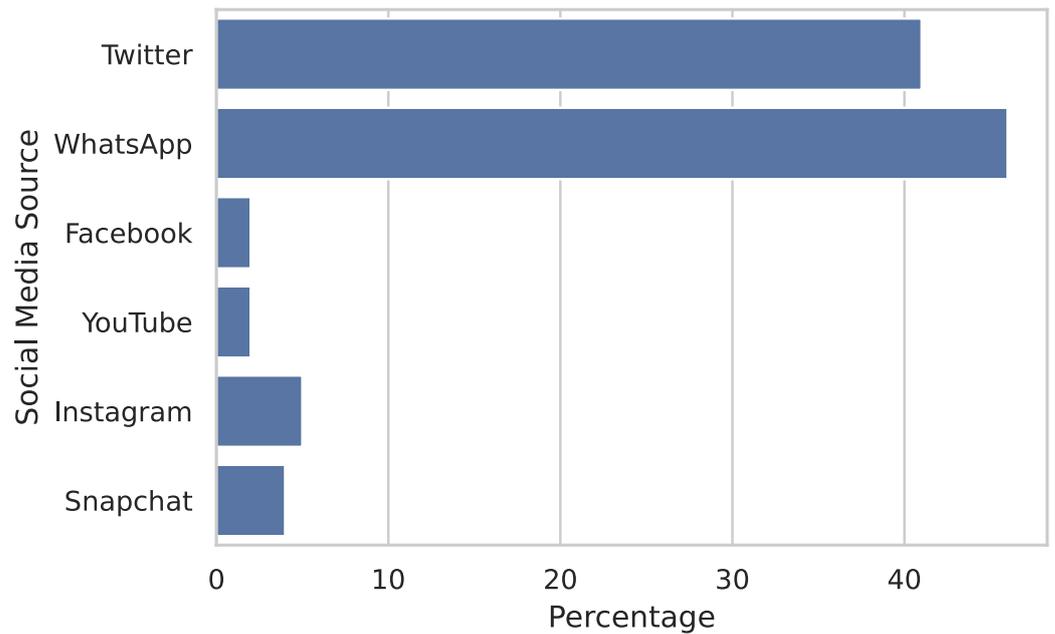
Figure 5. Social Media Sources of Misinformation Reported by the Community.

The content of the community-reported misinformation mainly covers seven different areas as shown in Figure 6, which include health-related advice, conspiracy theories, biological war, China and the source of COVID-19, local Saudi policies, 5G Networks, and Arab strong immunity to COVID-19. While the sample surveyed is small (88 participants), we see great overlap with the topics of misinformation identified in our analysis of Twitter data published during the collection period (i.e., December 2019–April 2020). Thus, we concluded that data saturation was reached with current sample size as no new information/themes were observed in the survey responses.
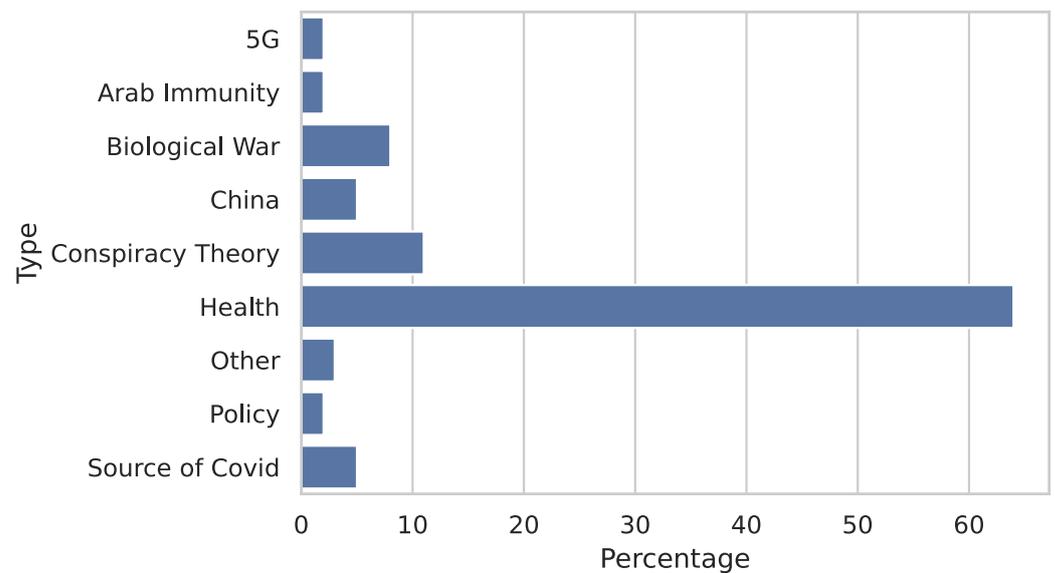
Figure 6. Types of reported misinformation.

5. Discussion

This study aimed to address the following research questions: (1) What forms of misinformation spread the most in pandemics? (2) How does misinformation evolve over time?

In this current research, a retrospective analysis of the COVID-19 infodemic was conducted concerning the country of Saudi Arabia. According to the three basic research questions, our major contributions can be summarized as follows: (1) We extracted a sample of tweets from a large Arabic dataset related to the COVID-19 pandemic from December 2019 to April 2020. Human annotators were utilized for labeling the sample for this purpose. (2) We utilized quantitative and qualitative methodologies including Twitter data and survey results to understand public opinion toward misinformation. (3) We discussed the findings through the lens of the local context in Saudi Arabia and looked at how misinformation is spread depending on the culture, laws, and period of time.

Regarding the first research question, the narrative threads of misinformation that have circulated in Saudi Arabia since the COVID-19 pandemic include the origin of COVID-19 and health advice coping with COVID-19. The findings suggest that misinformation could be tied in with sense-making, which is consistent with previous research [1,2]. It has been concluded that people turn to rumors as a way to cope with uncertainty. Moreover, the types of misinformation that were shared included health-related rumors and political issues, which could indicate that the public is more interested in these subjects and is willing to believe and share such information without validating it with an authentic source. In general, online platforms provide a venue for finding and sharing health information [34,35]. It is worth noting that the government's prompt response to the pandemic spread of disinformation has aided in limiting the transmission of misinformation to the public and reducing the duration of rumors among individuals. This consistent information was also reported by other research work for the African subcontinent [36].

In addressing the second research question, this paper identifies temporal patterns of misinformation. It can be observed from Figure 3 that multiple misinformation items appeared in the mid-set of the pandemic with a quick turnaround of misinformation as the perception of Islamophobia and conspiracy theories. Moreover, it can also be noticed from Figure 3 that there is no single pattern for how misinformation is shared in the community; however, it is evident that it takes a couple of rounds of sharing before it times out. Further examining the results, the weekly growth of misinformation demonstrated that the biological war exhibits the highest value of the thematic one compared to the others. Moreover, the other pandemic set showed quite similar responses in the time run; albeit, a little scatter was observed for Islamophobia, health advice, and conspiracy theories of thematic categories in the digital misinformation systems. Such responses in the results of misinformation growth might lead to the fact that the mean respondent has disseminated the most in social media via Tweets.

It was evident that misinformation presents a severe risk to public health and public action. This finding is in line with WHO's infodemic briefing stating, "Analysis of social networks have shown polarization for COVID-19 health topics, and this polarization is exacerbating information bottlenecks, making it difficult to ensure universal access to credible health information. Network analysis can also be enabled to identify influential users within a network, including how closely connected they are to other influential users in order to better understand the opinion of drivers on a specific issue. For COVID-19 case, social media data can help to characterize trends, the type of information spreading across platforms, and spread of information using epidemic models, as well as the diffusion of varying levels of (in)accurate information [19]." Heba et al. [37] studied the transmission of the COVID-19 pandemic in Saudi Arabia and found that taking preventive measures resulted in a 27 percent reduction in infection and death rates, which has a direct influence on public health and public action initiatives. This reduction of infection and death rate is mainly attributed to the fact that the misinformation is not widely spread over Twitter or actual information reaches to the actual correspondence.

Since the outbreak of COVID-19 in Saudi Arabia, the country has taken a number of actions to limit the spread of the actual virus as well as any misinformation related to the virus. The sort of actions included lockdowns in many private and public sectors and services [38]. Following these circumstances, Heba et al. [37] reported a similar form of government effort to prevent the virus's spread in their investigation. At the very early stages of the COVID-19 epidemic, the Saudi government implemented travel prohibitions for all countries, schools and universities were converted to distance learning, all international flights were postponed, and even the five daily prayers were outlawed throughout the country. Persons who spread rumors or false information on social media could face jail time up to five years, be fined up to a SR 1 million, or face both punishments [15].

In addition to issuing new laws, Saudi Arabia attempted to raise the public's awareness of the virus by disseminating information from reliable sources. An example of sharing updated information widely with the public is the daily news conference conducted by the Ministry of Health in Saudi Arabia. Further, Saudi Arabia dedicated a particular number (937) for people who want to learn more about the virus from trusted sources [39]. It is noticeable that mass awareness plays a vital role in assisting and maintaining government interferences and limits the spread of the virus with misinformation in the public platform. The greater awareness should focus on individuals like older people, and cultural minorities are represented as at high risk of COVID-19 in the country [40,41].

Moreover, many digital applications have been created during the pandemic to provide services related to COVID-19 to the people of Saudi Arabia. Examples of these applications are Tawakklna (https://ta.sdaia.gov.sa/en/index, accessed on 27 September 2021) and Sehhaty (https://www.moh.gov.sa/en/eServices/Pages/Vaccine-date.aspx, accessed on 27 September 2021), which provide health information and services for the people of Saudi Arabia [38]. Saudi Arabia issued these new policies and services to minimize the spread of the virus and any misinformation related to it. With the digital appliances, personal awareness of protective measures is the dominating factor in limiting the wide range of spreading the COVID-19 epidemic in any country [42].

Our investigation of misinformation in Saudi Arabia is strategically framed by previous work on misinformation and disinformation. Much of the early NLP work focused on trust, the credibility of Twitter content, and extremist narratives (e.g., [43,44]). The studies by Alshaalan et al. [6,8] suggest that social media has played a vital role in facilitating fear, anxiety, and hatred in politically charged and volatile environments. In the context of public health, designing evidence-based interventions to protect the public and mitigate misinformation during an infodemic relies heavily on robust and responsive automated methods. Along with the social media coverage, it is important to maintain a high index of critical indicators to combat the COVID-19 pandemic such as applying strict infection protocols, active surveillance measures, and attending mandatory online educational short courses about the current pandemic scenario in the country [45].

The urgency and rapid changes in the ongoing pandemic cause some limitations for the current study. This includes the data sampling technique, as the Twitter dataset only provides snapshots of the current public perceptions and psychological crisis responses and that will not allow the assessment of genuine causal relationships. Moreover, a significant amount of public perceptions are expressed and disseminated through encrypted platforms such as Whatsapp and private communication, which are beyond the scope of this study's analytics. Another drawback is that both textual and multimedia misinformation contain innuendo and nuance, which are difficult to quantify accurately with the current machine learning algorithm platform. However, it should be emphasized that the sampling method for the COVID-19 infodemic is currently based on a limited dataset, and more research is needed to find the best approaches for capturing the full spectrum of public responses.

Concerning further work on COVID-19 in this specific region, it is is suggested to expand current work by incorporating varying machine learning model misinformation themes. Furthermore, we look forward to determining the impact of governmental laws on

the dissemination of misinformation on social media and its related risk factors. We also look forward to utilizing the machine learning classifiers on our initial annotated dataset.

6. Conclusions

This work describes a technical approach to social media analytics that aims to strengthen health systems by detecting emerging and resurgent health threats in the form of misinformation or disinformation. The number of COVID-19 infections continued to increase after the first infected patients were found in Saudi Arabia. The most popular platform for spreading pandemic misinformation is social media and community digitization. Developing social media listening approaches for teams to detect changes in public discourse or narratives during pandemics contributes to creating a more adaptive and effective public health emergency response.

This study demonstrates that the social media platform plays a critical role in disseminating disinformation in the public sphere. It is the most significant source of rumors, particularly misinformation about pharmaceutical corporations. It also suggests that precautionary measures such as ignoring the misinformation, appropriate methods of using technology, government legislation, distance learning, remote working, and social and self-awareness might significantly limit the spread of the pandemic. The country's government should pay special attention to what steps are being made to prevent misinformation from spreading through internet platforms, as they are important platforms in virus transmission.

Appendix A

Table A1 lists hashtags populated by Saudi governmental Twitter accounts. These hashtags urge the community to be responsible about decreasing the number of cases by following prevention measures, reassure the community about the availability of products, and answer common questions about COVID-19. In addition, hashtags that mainly discuss precautionary measures governments have applied. These include discussions of curfew, business closures, and travel restrictions. The table shows the list of hashtags in Arabic accompanied by English translation.

Table A1. Lists of popular hashtags used by the public in Saudi Arabia along with English translation.

Hashtags	English Translation
الوقاية _من _كورونا	Corona Prevention
كلنا _مسؤول	We are all responsible
عش _بصحة	Live healthily
أسئلة _كورونا	Corona's Questions
أبطال _الصحة	Health Heroes
أبطال _المجتمع	Community Heroes
المنتجات _متوفرة	Products available
الخدمات _مستمرة	Services continuous
متر _ونص	One and a half meters
شكراً _أبطال _التعليم	Thanks Education heroes
إيقاف _صلاة _الجماعة	Stopping congregational prayer
إغلاق _الحدائق	nParks closure
صلوا _في _رحالكم	Pray in your travel
ايقاف _الصلاة _بالمسجد	Stop praying in the mosque
ايقاف _صلاة _الجمعة _والجماعة	Stopping Friday and group prayers
إغلاق _محلات _الحلاقة	barber shops closure
اغلاق _المقاهي	Cafes closure
اغلاق _الصالونات	Salons closure
ايقاف _الدوري	Stopping football league
تعليق _النشاط _الرياضي	Sports suspension
تعليق _الرحلات _الدوليه	International flights suspended
تعليق _الرحلات _الداخليه	Internal flights suspended
تعليق _العمل	Work suspension
تعليق _الدراسة	School Suspension
اغلاق _النوادي _الرياضية	Gyms closure
اغلاق _المولات _في _السعوديه	Closure of the malls in Saudi Arabia
إغلاق _المجمعات _التجارية	Closure of Shopping Centres
تعليق _القطاع _الخاص	private sector suspension
منع _التجول	Curfew
منع _التنقل _بين _المناطق	Prevent movement between regions

References

1. Akbar, S.Z.; Panda, A.; Kukreti, D.; Meena, A.; Pal, J. Misinformation as a Window into Prejudice: COVID-19 and the Information Environment in India. *Proc. ACM Hum.-Comput. Interact.* **2021**, *4*, 249. [CrossRef]
2. Gallotti, R.; Valle, F.; Castaldo, N.; Sacco, P.; De Domenico, M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. *Nat. Hum. Behav.* **2020**, *4*, 1285–1293. [CrossRef]
3. Gmi_blogger. Saudi Arabia Social Media Statistics 2020 (Infographics)—GMI Blog. 2020. Available online: https://froggyads.com/blog/saudi-arabia-social-media-statistics-infographics-gmi-blog/ (accessed on 27 September 2021).
4. Al-Masoudi, M. Twitter in Saudi Arabia. 2021. Available online: https://twitter.com/saudiarabia (accessed on 27 September 2021).

5. Alamro, N.; Almana, L.; Alabduljabbar, A.; AlKahtani, M.; AlDihan, R.; Almansour, A.; Alobaid, N.; AlOthaim, N.; Alshunaifi, A. Saudi Arabia COVID-19 Snapshot MOnitoring (COSMO Saudi): Monitoring Knowledge, Risk Perceptions, Preventive Behaviours, and Public Trust in the Current Coronavirus Outbreak in Saudi Arabia. PsychArchives. 2020. Available online: https://www.psycharchives.org/handle/20.500.12034/2496 (accessed on 27 September 2021).
6. Alharbi, A.; Lee, M. Kawarith: An Arabic Twitter Corpus for Crisis Events. In Proceedings of the Sixth Arabic Natural Language Processing Workshop, online, 19–23 April 2021; Association for Computational Linguistics: Kyiv, Ukraine, 2021; pp. 42–52.
7. Alqurashi, S.; Alhindi, A.; Alanazi, E. Large Arabic Twitter Dataset on COVID-19. *arXiv* **2020**, arXiv: 2004.04315.
8. Alshalan, R.; Al-Khalifa, H.; Alsaeed, D.; Al-Baity, H.; Alshalan, S. Detection of Hate Speech in COVID-19—Related Tweets in the Arab Region: Deep Learning and Topic Modeling Approach. *J. Med. Internet Res.* **2020**, *22*, e22609. [CrossRef] [PubMed]
9. Alsudias, L.; Rayson, P. COVID-19 and Arabic Twitter: How can Arab World Governments and Public Health Organizations Learn from Social Media? In Proceedings of the NLP COVID-19 Workshop, Seattle, WA, USA, 9 July 2020; Association for Computational Linguistics: Kyiv, Ukraine, 2020.
10. Oyeyemi, S.O.; Gabarron, E.; Wynn, R. Ebola, Twitter, and misinformation: A dangerous combination? *BMJ* **2014**, *349*, g6178. [CrossRef]
11. Chew, C.; Eysenbach, G. Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak. *PLoS ONE* **2010**, *5*, e14118. [CrossRef]
12. Badawy, A.; Ferrara, E.; Lerman, K. Analyzing the Digital Traces of Political Manipulation: The 2016 Russian Interference Twitter Campaign. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, Spain, 28–31 August 2018; pp. 258–265. [CrossRef]
13. Ferrara, E. Manipulation and abuse on social media. *ACM SIGWEB Newsl.* **2015**, 1–9. [CrossRef]
14. Depoux, A.; Martin, S.; Karafillakis, E.; Preet, R.; Wilder-Smith, A.; Larson, H. The pandemic of social media panic travels faster than the COVID-19 outbreak. *J. Travel Med.* **2020**, *27*, taaa031. [CrossRef]
15. SR1m Fine and 5 years Jail for Violating Coronavirus Measures—Saudi Gazette. 2020. Available online: https://saudigazette.com.sa/article/592724/SAUDI-ARABIA/SR1m-fine-and-5-years-jail-for-violating-coronavirus-measures (accessed on 27 September 2021).
16. Haouari, F.; Hasanain, M.; Suwaileh, R.; Elsayed, T. ArCOV19-Rumors: Arabic COVID-19 Twitter Dataset for Misinformation Detection. *arXiv* **2021**, arXiv:2010.08768.
17. Jussila, J.; Suominen, A.; Partanen, A.; Honkanen, T. Text analysis methods for misinformation-related research on Finnish language twitter. *Future Internet* **2021**, *13*, 157. [CrossRef]
18. Alqurashi, S.; Hamoui, B.; Alashaikh, A.; Alhindi, A.; Alanazi, E. Eating Garlic Prevents COVID-19 Infection: Detecting Misinformation on the Arabic Content of Twitter. *arXiv* **2021**, arXiv:2101.05626.
19. WHO. *4th Virtual WHO Infodemic Management Conference: Advances in Social Listening for Public Health*; WHO: Geneva, Switzerland, 2021.
20. Adel, G.; Wang, Y. Arabic Twitter Corpus for Crisis Response Messages Classification. In Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence, Sanya, China, 20–22 December 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 498–503. [CrossRef]
21. Haouari, F.; Hasanain, M.; Suwaileh, R.; Elsayed, T. ArCOV-19: The First Arabic COVID-19 Twitter Dataset with Propagation Networks. In Proceedings of the Sixth Arabic Natural Language Processing Workshop, online, 19–23 April 2021; Association for Computational Linguistics: Kyiv, Ukraine, 2021; pp. 82–91.
22. Mubarak, H.; Hassan, S. ArCorona: Analyzing Arabic Tweets in the Early Days of Coronavirus (COVID-19) Pandemic. *arXiv* **2021**, arXiv:2012.01462.
23. Addawood, A. Coronavirus: Public Arabic Twitter Data Set. 2020. Available online: https://www.preprints.org/manuscript/202004.0263/v1 (accessed on 27 September 2021).
24. Alam, F.; Shaar, S.; Dalvi, F.; Sajjad, H.; Nikolov, A.; Mubarak, H.; Martino, G.D.S.; Abdelali, A.; Durrani, N.; Darwish, K.; et al. Fighting the COVID-19 Infodemic: Modeling the Perspective of Journalists, Fact-Checkers, Social Media Platforms, Policy Makers, and the Society. *arXiv* **2020**, arXiv:2005.00033.
25. Tangcharoensathien, V.; Calleja, N.; Nguyen, T.; Purnat, T.; D'Agostino, M.; Garcia-Saiso, S.; Landry, M.; Rashidian, A.; Hamilton, C.; AbdAllah, A.; et al. Framework for managing the COVID-19 infodemic: Methods and results of an online, crowdsourced WHO technical consultation. *J. Med. Internet Res.* **2020**, *22*, e19659. [CrossRef]
26. UNESCO. *Combating the Disinfodemic: Working for Truth in the Time of COVID-19;* United Nations Educational, Scientific and Cultural Organization: Paris, France, 2020.
27. News, U. *During this Coronavirus Pandemic, 'Fake News' is Putting Lives at Risk*; UNESCO: Paris, France, 2020.
28. Murphy, J. International perspectives and initiatives. *Health Inf. Libr. J.* **2007**, *24*, 62–68. [CrossRef]
29. Chen, E.; Jiang, J.; Chang, H.-C.H.; Muric, G.; Ferrara, E. COVID-19 Infodemiology at Planetary Scale: Charting the Information and Misinformation Landscape to Characterize Misinfodemics Spread on Social Media. *JMIR Prepr.* **2021**.
30. Vargas, L.; Emami, P.; Traynor, P. On the Detection of Disinformation Campaign Activity with Network Analysis. In Proceedings of the 2020 ACM SIGSAC Conference on Cloud Computing Security Workshop, New York, NY, USA, 9 November 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 133–146. [CrossRef]
31. Venkatesh, V.; Brown, S.A.; Bala, H. Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS Q.* **2013**, *37*, 21–54. [CrossRef]

32. Biernacki, P.; Waldorf, D. Snowball sampling: Problems and techniques of chain referral sampling. *Sociol. Methods Res.* **1981**, *10*, 141–163. [CrossRef]
33. Al-Zaman, M.S. A Thematic Analysis of Misinformation in India during the COVID-19 Pandemic. *Int. Inf. Libr. Rev.* **2021**, [CrossRef]
34. Alasmari, A.; Zhou, L. Share to Seek: The Effects of Disease Complexity on Health Information Seeking Behavior. *J. Med. Internet Res.* **2021**, *23*, e21642. [CrossRef]
35. Alasmari, A.; Zhou, L. How multimorbid health information consumers interact in an online community Q&A platform. *Int. J. Med. Inform.* **2019**, *131*, 103958. [CrossRef]
36. Ahinkorah, B.; Ameyaw, E.; Hagan, J., Jr.; Seidu, A.A.; Schack, T. Rising Above Misinformation or Fake News in Africa: Another Strategy to Control COVID-19 Spread. *Front. Commun.* **2020**, *5*, 45. [CrossRef]
37. Adly, H.; Aljahdali, I.; Garout, M.; Khafagy, A.; Saati, A.; Saleh, S. Correlation of COVID-19 Pandemic with Healthcare System Response and Prevention Measures in Saudi Arabia. *Int. J. Environ. Res. Public Health* **2020**, *17*, 6666. [CrossRef]
38. Hassounah, M.; Raheel, H.; Alhefzi, M. Digital Response During the COVID-19 Pandemic in Saudi Arabia. *J. Med. Internet Res.* **2020**, *22*, e19338. [CrossRef] [PubMed]
39. Saudis Fight Misinformation Related to Coronavirus Disease | Arab News. 2020. Available online: https://www.arabnews.com/node/1662936/saudi-arabia (accessed on 27 September 2021).
40. Samir Abdelhafiz, A.; Mohammed, Z.; Ibrahim, M.; Ziady, H.; Alorabi, M.; Ayyad, M.; Sultan, E. Knowledge, Perceptions, and Attitude of Egyptians Towards the Novel Coronavirus Disease (COVID-19). *J. Community Health* **2020**, *45*, 881–890. [CrossRef]
41. Wolf, M.; Serper, M.; Opsasnick, L.A.; O'Conor, R.; Curtis, L.; Benavente, J.Y.; Wismer, G.; Batio, S.; Eifler, M.; Zheng, P.; et al. Awareness, Attitudes, and Actions Related to COVID-19 Among Adults With Chronic Conditions at the Onset of the U.S. Outbreak. *Ann. Intern. Med.* **2020**, *173*, 100–109. [CrossRef] [PubMed]
42. Zhang, L.; Shen, M.; Ma, X.; Su, S.; Gong, W.; Wang, J.; Tao, Y.; Zou, Z.; Zhao, R.; Lau, J.; et al. What Is Required to Prevent a Second Major Outbreak of SARS-CoV-2 upon Lifting Quarantine in Wuhan City, China. *Innovation* **2020**, *1*, 100006. [CrossRef]
43. Nouh, M.; Nurse, J.R.; Goldsmith, M. Understanding the Radical Mind: Identifying Signals to Detect Extremist Content on Twitter. In Proceedings of the 2019 IEEE International Conference on Intelligence and Security Informatics (ISI), Shenzhen, China, 1–3 July 2019; pp. 98–103. [CrossRef]
44. Al-Twairesh, N.; Al-Khalifa, H.; Al-Salman, A. Towards analyzing Saudi tweets. In Proceedings of the 2015 First International Conference on Arabic Computational Linguistics (ACLing), Cairo, Egypt, 17–20 April 2015; pp. 114–117.
45. WHO. *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*; WHO: Geneva, Switzerland, 2020.