ISSN 2228-9860 eISSN 1906-9642 CODEN: ITJEA8

International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies

http://TuEngr.com

Current Trend of Virology in the Post COVID-19 Era with Related Topic-Based Tweet Analysis

Fahim K Sufi[1*], Musleh AlSulami[2*]

[1] *Federal Government, Melbourne, AUSTRALIA.*

[2] *Umm Al-Qura University, Makkah, Makkah, SAUDI ARABIA.*

**Corresponding Author (Email: research@ fahimsufi.com, mhsulami@ uqu.edu.sa)*

Paper ID: 13A4T

Volume 13 Issue 4

Received 23 December 2021
Received in revised form 19 March 2022
Accepted 06 April 2022
Available online 14 April 2022

Keywords:
Social media analysis; Tweet analysis; Covid tweet; Tweeter hashtag; Covid hashtag; number of posts; number of users; Engagement metric; Influence metric; Word cloud on Virology; Virological Science; Preventive vaccines; Therapeutic vaccines; Viral pathogenesis; Viral immunology; Antiviral strategies; Virus structure; Virus expression; Viral resistance; Emerging viruses; Interspecies transmission; Viruses and cancer; Viral diseases.

Abstract

In light of current trends in virology, we performed social media analysis of 13 main topics in the area of virology and ranked these topics with metrics such as users, posts, engagement, and influence. These metrics were monitored against the 13 keywords on Twitter for the same period (i.e., from 27 November to 6 December 2021) for benchmarking purposes. The 13 main topics were "virological Science", "preventive vaccines", "therapeutic vaccines", "viral pathogenesis", "viral immunology", "antiviral strategies", "virus structure", "virus expression", "viral resistance", "emerging viruses", "interspecies transmission", "viruses and cancer" and "viral diseases". "viral diseases" recorded the highest number of users (i.e., 905 users) and the highest number of post (i.e., about 1K posts). The second-highest number of posts were monitored to be on "therapeutic vaccines" with 729 posts from 691 users. In terms of engagement, "viral diseases" (3.4 K) were found to be on the top followed by "viruses and cancer" (3.1K). Lastly, in terms of influence, "viral diseases" recorded 9.0 million influences followed by 6.6 million influences on "emerging viruses". In summary, "viral diseases" was found to be the most engaging and influential topic highest with the highest number of posts from most of the tweet users. In relation to trending hashtags in virology, #COVID19 recorded the highest number of hashtags, followed by #omicron, #sarscov2, #publichealth, #omicronvarient, #wuhan, #originofcovid, #fauci and #epidemiology. Word clouds showing the main area of discussion were also generated for these 13 main topics.

Disciplinary: Sociology, Information Technology (Social Media and Networking).

Cite This Article:

Sufi, F. K., AlSulami, M. (2022). Current Trend of Virology in the Post COVID-19 Era with Related Topic-Based Tweet Analysis. *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies, 13*(4), 13A4T, 1-10. http://TUENGR.COM/V13/13A4T.pdf DOI: 10.14456/ITJEMAST.2022.83

1 Introduction

In the era of post-COVID-19, more people are interested in virology and associated topics. People are constantly posting social media updates on topics related to virology. On the other hand, several other social media followers are being engaged with these posts by liking these posts, commenting on these posts or even sharing these posts. Hence, by using modern techniques in social media analysis, we can easily identify trending topics of discussion on social media [1] [2].

Unlike our previous research in knowledge discovery on landslides [3] [4], abnormality detection [5] [6] and person identification [7], where Machine Learning (ML), and Artificial Intelligence (AI) techniques were used, this paper does not use any ML or AI-based techniques. Studies in [3] [4] [5] [6] [7] relied on feature extraction and applied AI and ML-based techniques to these extracted features. However, in this paper, we used simple statistical techniques without the use of AI (similar to [8]). Hence, the methodology described in this paper is highly suitable for a social media analyst or researcher who is not familiar with AI or ML, or data science.

Figure 1: The 13 related topics in Virology

In this study, firstly, we investigated the topic of Virology and created a list of 13 related topics. These related topics were "Virological Science", "Preventive vaccines", "Therapeutic vaccines", "Viral pathogenesis", "Viral immunology", "Antiviral strategies", "Virus structure", "Virus expression", "Viral resistance", "Emerging viruses", "Interspecies transmission", "Viruses and cancer" and "Viral diseases". Figure 1 shows these related sub-topics. Then, for each of these

sub-topics, we obtained 4 different metrics for measuring the trends. These metrics were the number of posts, number of users, engagement, and influence, further defined in Methodology.

In the results and discussion section, we show that "Viral diseases" recorded the highest number of users (i.e., 905 users) and the highest number of posts (i.e., about 1K posts). In terms of engagement, Viral diseases Viruses, and cancer have the highest level of engaengagement4 thousand and 3.1 thousand respectively). Viral diseases had the highest number of influences (9 million). Moreover, word clouds and hashtag rankings for each of these 13 sub-topics are also depicted within the results and discussion section. Omicron was the most popular phrase under the emerging virus sub-topic.

Using the technique and metrics demonstrated in this study, any social media researcher can obtain insights into any trending topics.

2 Methodology

Using Tweeter Application Programming Interfaces (APIs), a researcher can obtain valuable metrics like the number of posts on a particular topic, number of users posting on that topic, number of likes on posts, number of comments on posts, number of retweets or shares, number of followers for users, number of friends for users and various other metrics. Metrics like the number of likes on posts, the number of comments on posts, and a number of retweets or shares indicate a composite metric called *Engagement*. On the other hand, a number of followers for users and the number of friends for users indicates another composite metric called *Influence*. Other than using direct APIs, social media researchers can also obtain these statistics via a number of third-party tools like Microsoft Power Automate, Microsoft Power Query and others as depicted in our previous research [1] [2].

This research uses APIs developed in [1] [2], to obtain four metrics: 1) Number of Posts per topic, 2) Number of Users posting per topic, 3) Number of Engagement (i.e., a combination of likes, comments and shares), 4) Number of Influence (i.e., users followers and friends) on all relevant sub-topics using Algorithm 1. Firstly, all the relevant sub-topics are generated. Secondly, the 4 metrics (i.e., number of posts, number of users, engagement and influence) are obtained using API's and News sensors developed in our earlier research [1] [2]. It should be noted that these metrics were obtained for all the 13 related topics in Virology as depicted in Figure 1.

Algorithm 1: Related Topic-based Tweet Analysis

Create a list of all the relevant topics
For Each Topic in the list of Related Topics
    Get the Number of Posts
    Get Number of Users
    Get Engagement Statistics
    Get Influence
Next

3 Results & Discussion:

Using Algorithm 1, we obtained four metrics (i.e., Users, Posts, Engagement and Influence) for all the 13 related topics in Virology. Table 1 shows the summarized result.

Table 1: Users, posts, engagement and influence for all the 13 topics in Virology

Keywords Analyzed	Users	Posts	Engagement	Influence
Virology	1.8K	2.1K	7.9K	12.3M
Virological Science	16	16	31	29.2K
Preventive vaccines	227	247	693	6.3M
Therapeutic vaccines	691	729	2.8K	2.5M
Viral pathogenesis	30	30	83	78.4K
Viral immunology	125	146	419	977.1K
Antiviral strategies	8	8	28	8.0K
Virus structure	492	519	1.3K	3.0M
Virus expression	45	50	100	392.1K
Viral resistance	498	641	1.9K	2.4M
Emerging viruses	224	233	570	6.6M
Interspecies transmission	5	6	3	1.1K
Viruses and cancer	665	690	3.1K	2.0M
Viral diseases	905	1.0K	3.4K	9.0M

As we can see from Table 1, 905 users posted on the *Viral diseases* sub-topic. Therefore, the highest number of users posted on *Viral diseases* followed by *Viruses and cancer*. The viral diseases sub-topic also accumulated the highest number of posts (i.e., approximately, 1 thousand posts). With 729 posts, the *Therapeutic vaccines* sub-topic received the second-highest number of posts. In terms of engagement, *Viral diseases Viruses and cancer* had the highest level of engagement (3.4 thousand and 3.1 thousand respectively). As mentioned in the Methodology section, engagement refers to the likes, comments and shares of users on a particular topic. *Viral diseases* had the highest number of influences (approximately 9 million). The second highest influence was noted on the topic "*Emerging viruses*" with 6.6 million influence score. As mentioned before, influence score refers to the total number of friends or followers for the users posting on particular topics.

Word cloud images show the frequencies of key phrases used within a topic. For example, Figure 2 shows the word cloud for Virology. As we can see from Figure 2, Lab, Institute, Variant, viruses, omicron, Wuhan, facilities are the highest occurring phrases within the topic of Virology.

Figure 2: Word cloud on Virology.

Table 2 shows the word clouds for each of the 13 sub-topics of Virology. Figure 3 shows the ranking of hashtags for Virology. After the hashtag virology, COVID-19 and Omicron dominated as

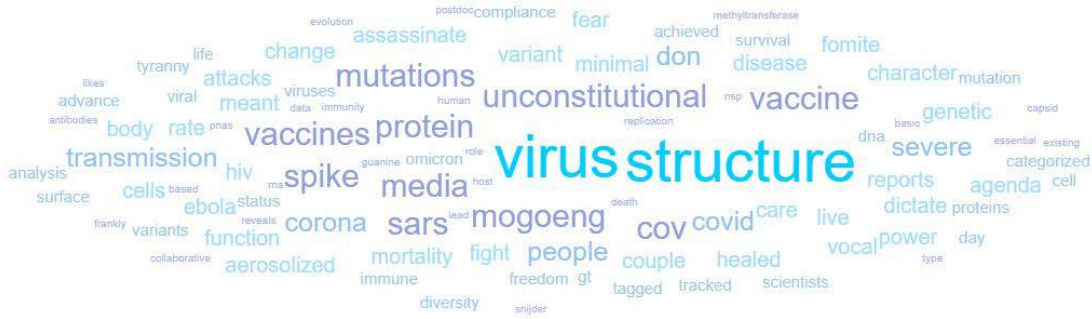
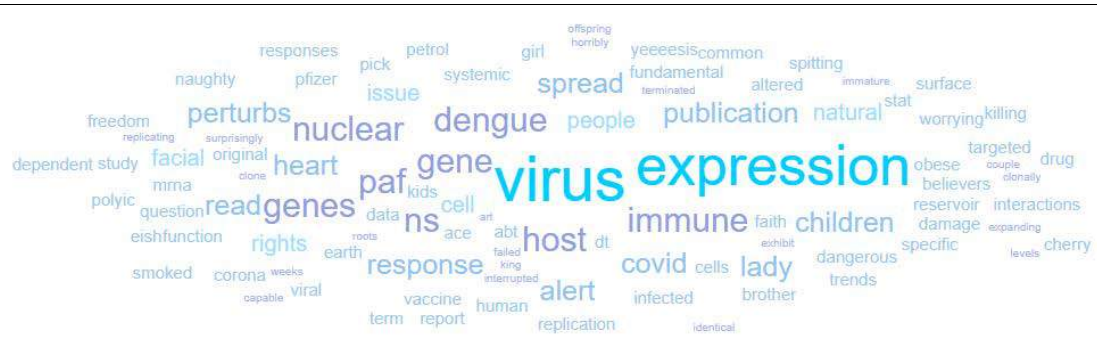
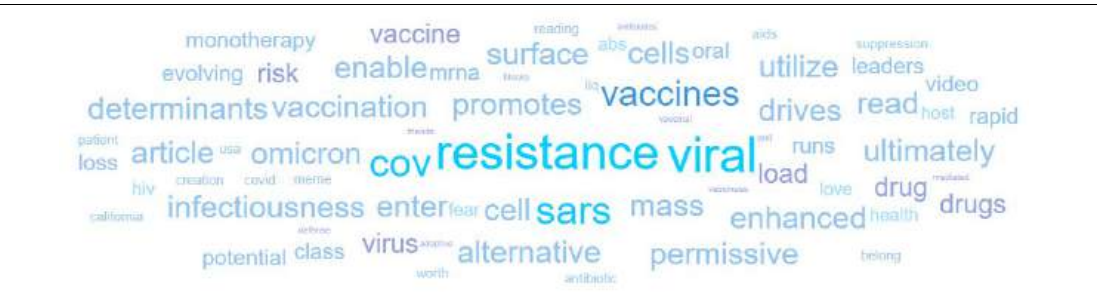
the most popular hashtags on the topic of virology. Table 3 shows the hashtag frequencies for the 13 sub topics.

Hashtags	
#virology	42
#covid19	39
#omicron	25
#sarscov2	23
#publichealth	14
#omicronvarient	13
#wuhan	13
#originofcovid	12
#fauci	12
#epidemiology	11

Figure 3: Hashtags count in Virology.

Table 2: Word Cloud showing the key phrases for each of the 13 sub-topics

	Keywords / Topic with the main points of discussion in Word Cloud
Virological Science	
Preventive vaccines	
therapeutic vaccines	
Viral pathogenesis	

Keywords / Topic with the main points of discussion in Word Cloud	
Viral immunology	
Antiviral strategies	
Virus structure	
Virus expression	
viral resistance	
Emerging viruses	

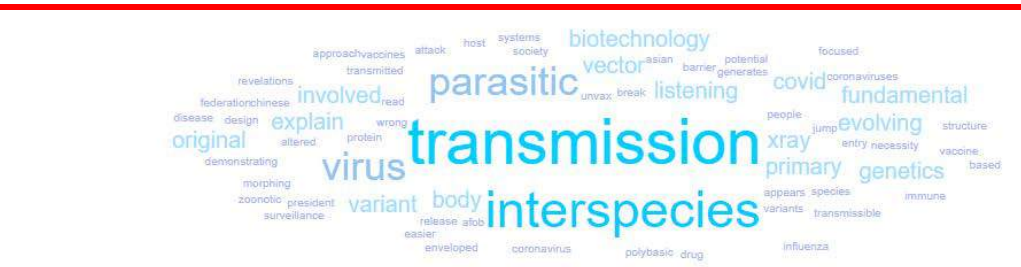
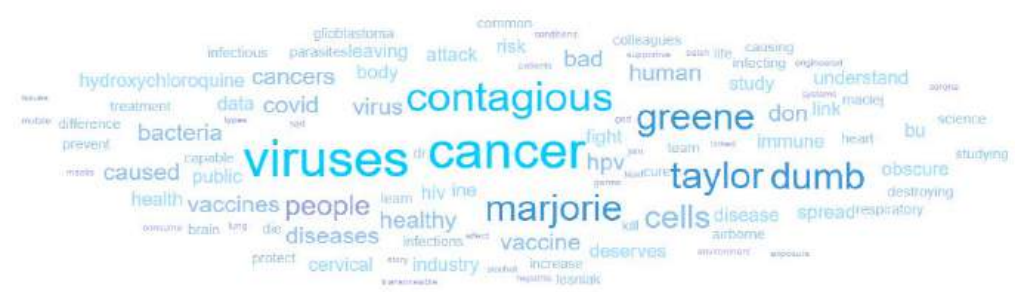
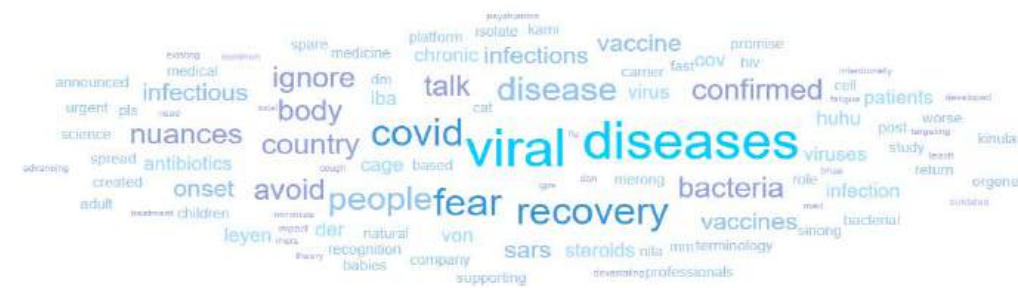
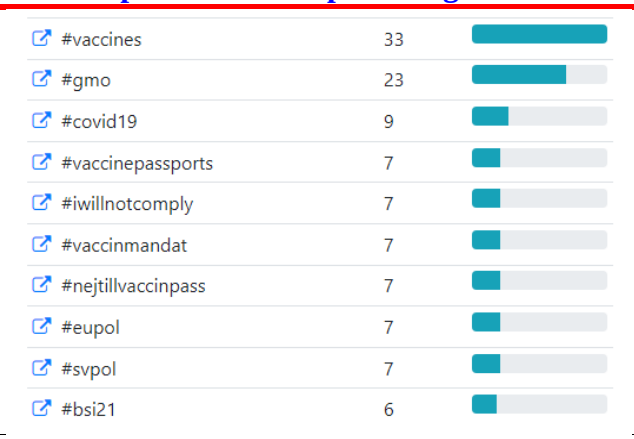
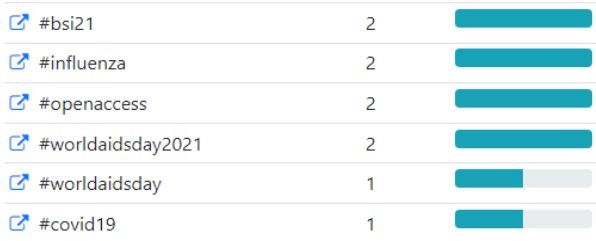
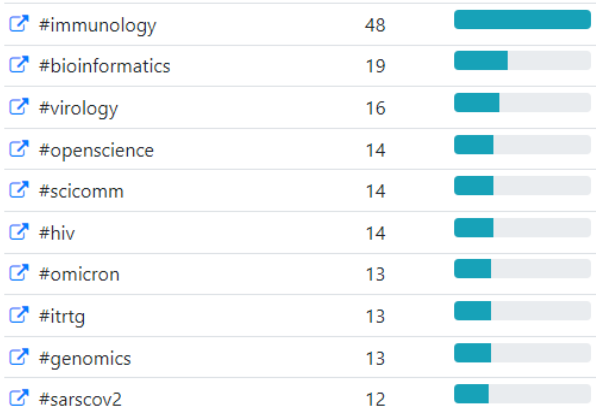
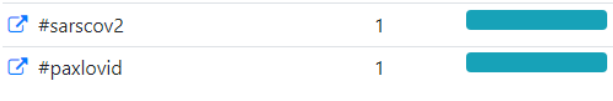
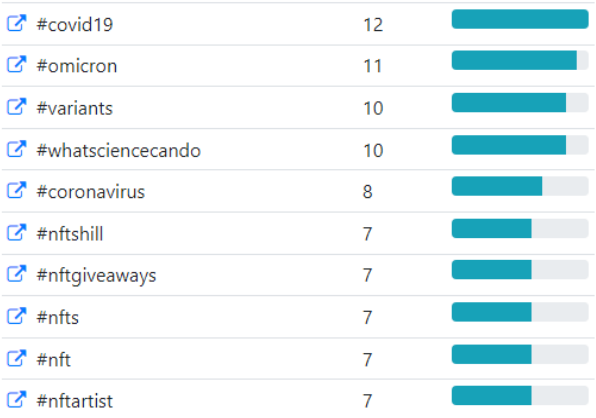
Keywords / Topic with the main points of discussion in Word Cloud	
Interspecies transmission	
Viruses and cancer	
viral diseases	

Table 3: Frequencies of hashtags showing trends for each of the 13 sub-topics

Keywords / Topics with the top hashtags		
Virological Science	#openscience	1
	#scholcomm	1
	#abstract	1
	#antibodies	1
	#sabin	1
	#pseudoviruses	1
	#endpolio	1
	#polio	1
	#vaccineswork	1
	#publichealth	1
Preventive vaccines	#covid19	8
	#flushot	6
	#flu	6
	#aidsday	4
	#aidsawareness	4
	#aids	4
	#allemand	3
	#omicron	3
	#rd	2
	#vaccines	2

Keywords / Topics with the top hashtags	
therapeutic vaccines	
Viral pathogenesis	
Viral immunology	
Antiviral strategies	
Virus structure	
Virus expression	

Keywords / Topics with the top hashtags	
viral resistance	#novaccinemandates 80 #covidvaccine 80 #novaccinepassport 80 #vaccinesideeffects 80 #omicron 68 #covid19 67 #quercetin 20 #nubtc 18 #myocarditis 15 #cytokinestorm 10
Emerging viruses	#omicron 40 #popab 14 #covid19 12 #lovevirology 8 #viruses 7 #sarscov2 6 #pandemics 6 #climatechange 4 #nervtag 4 #hiv 3
Interspecies transmission	#recombinant 1 #rodent 1 #cleavagesite 1 #onehealth 1
Viruses and cancer	#cancer 26 #covid19 18 #hpv 12 #alcoholawareness 10 #carcinogenic 10 #none 10 #galelaimmunebooster 8 #microbiome 3 #endometrialcancer 3 #healthconsumers 3
viral diseases	#longcovid 94 #mecfs 31 #covid19 23 #vaccineswork 19 #pwme 14 #increasetobaccotax 11 #coincidence 9 #watertech 8 #virology 7 #viraldiseasesmythologicalconceptions 7

4 Conclusion

From this study (Table 3), omicron is the most popular key phrase under sub-topic emerging viruses. Moreover, under the sub-topic *virus structure*, omicron is the second most discussed hashtag. Using the technique, described in this paper, a researcher can identify the most popular, influential or engaging topics or key phrases in social media.

In our previous research [3,4,5,6,7], ML and AI algorithms were used on various feature attributes for obtaining deeper insights. However, within this study ML and AI based techniques were not used. In future, we want to use AI based techniques on features extracted with Algorithm 1 (i.e., number of posts, users, engagement and influence). Using AI based techniques on these metric will give us deeper insights into social media analysis in future.

5 Availability of Data and Material

Data can be made available by contacting the corresponding author.

6 References

[1] F. K. Sufi and M. Alsulami, "Automated Multidimensional Analysis of Global Events With Entity Detection, Sentiment Analysis and Anomaly Detection," *IEEE Access*, vol. 9, pp. 152449 - 152460, 2021.

[2] F. K. Sufi, "AI-GlobalEvents: A Software for analyzing, identifying and explaining global events with Artificial Intelligence," *Software Impacts*, vol. Accepted in Press, no. DOI:10.1016/j.simpa.2022.100218, 2022.

[3] F. K. Sufi and M. Alsulami, "Knowledge Discovery of Global Landslides Using Automated Machine Learning Algorithms," *IEEE Access*, vol. 9, pp. 131400 - 131419, 2021.

[4] F. K. Sufi, "AI-Landslide: Software for acquiring hidden insights from global landslide data using Artificial Intelligence," *Software Impacts*, vol. 10, no. 100177, 2021.

[5] F. Sufi and I. Khalil, "Diagnosis of cardiovascular abnormalities from compressed ECG: a data mining-based approach," *IEEE transactions on information technology in biomedicine*, vol. 15, no. 1, pp. 33-39, 2010.

[6] F. Sufi and I. Khalil, "A clustering based system for instant detection of cardiac abnormalities from compressed ECG," *Expert Systems with Applications*, vol. 38, no. 5, pp. 4705-4713, 2011.

[7] F. Sufi and I. Khalil, "Faster person identification using compressed ECG in time critical wireless telecardiology applications," *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 282-293, 2011.

[8] F. Sufi, Q. Fang, I. Khalil and S. S. Mahmoud, "Novel methods of faster cardiovascular diagnosis in wireless telecardiology," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 4, pp. 537-552, 2009.

Dr.Fahim Sufi is a senior artificial intelligence solution architect with the federal government. He has held lead solution architect roles in several federal and state government agencies, including the Australian Department of Defence, the Australian Institute of Family Studies, the Victorian Department of Health, and the Victorian Department of Human Services. He obtained his PhD in computer science and information technology as well as a Master of Engineering in Computer Systems from RMIT University, Australia. His research interests include Artificial Intelligence, Machine Learning, Software Development, Big Data Analysis, Cyber, and Encryption.

Dr. Musleh Alsulami is an Assistant Professor of Information Systems at Umm Al-Qura University. He got a BSc in Software Engineering at Imam University, KSA, an MSc in Information Technology at Monash University, Australia, and a PhD in information systems at Monash University, Australia. His research interests include Enterprise Resources Planning (ERP), including ERP Life Cycles, Implementation Conflicts, Stakeholders, and Cloud-based ERP, as well as Digital Transformation in Government Organizations, Software Quality, and Human-Computer Interactions