2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Diachronic Analysis of Users' Stances on COVID-19 Vaccination in Japan using Twitter

Shohei Hisamitsu*† Sho Cho*
The University of Tokyo, Japan
{hisamits,cs}@tkl.iis.u-tokyo.ac.jp

Hongshan Jin Masashi Toyoda Naoki Yoshinaga
Institute of Industrial Science, the University of Tokyo, Japan
{jhs,toyoda,ynaga}@tkl.iis.u-tokyo.ac.jp

***Abstract*—To prevent and curb viral outbreaks, such as COVID-19, it is important to increase vaccination coverage while resolving vaccine hesitancy and refusal. To understand why COVID-19 vaccination coverage had rapidly increased in Japan, we analyzed Twitter posts (tweets) to track the evolution of people's stance on vaccination and clarify the factors of why people decide to vaccinate. We collected all Japanese tweets related to vaccines over a five-month period and classified the vaccination stances of users who posted those tweets by using a deep neural network we designed. Examining diachronic changes in the users' stances on this large-scale vaccine dataset, we found that a certain number of neutral users changed to a pro-vaccine stance while very few changed to an anti-vaccine stance in Japan. Investigation of their information-sharing behaviors revealed what types of users and external sites were referred to when they changed their stances. These findings will help increase coverage of booster doses and future vaccinations.**

I. Introduction

Vaccination is one of the most effective measures to prevent infectious diseases, such as measles and influenza. When a pandemic occurs, high vaccination coverage is required to end the pandemic and reduce deaths. Against the recent pandemic of COVID-19, delays in vaccination have become a major concern in many countries [1].

In Japan, which had ranked among the countries with the lowest vaccine confidence in the world [2], there was a great concern on vaccine uptake, especially among young people. According to a national survey on the intent to vaccinate against COVID-19 in February 2021, those who answered "unsure" and "no" were 32.9 and 11.0%, respectively [3]. However, after vaccination started on February 17, 2021, full vaccination coverage in Japan rapidly increased from June (3.4%) to October (75.3%) in 2021 and ranked eighth in the world.[1] The vaccination campaign thus progressed smoothly in Japan; those who had initially hesitated or refused to be vaccinated ended up being vaccinated.

Following studies on vaccine hesitancy and vaccine refusal using questionnaires and surveys [1]–[4], subsequent studies leveraged social media posts and focused on the COVID-19 vaccination [5]–[10]. Some of these studies revealed common reasons for vaccine hesitancy: concerns on the vaccine safety [5] and distrust in vaccine efficacy [6]. The other studies investigated people's attitudes towards vaccination: cross-country variations in positive and negative views on vaccination [7], spatiotemporal changes in sentiment on vaccines in the US [9], and the polarization between different vaccination communities in Japan [10]. Although these studies help us understand the current intent to vaccinate, none provide insights into diachronic changes in vaccine uptake.

To learn from Japan's successful case of increasing COVID-19 vaccine coverage, we analyzed the transition of users' vaccination intention on Twitter to reveal the factors that affected changes in their stances. By collecting all Japanese tweets related to vaccine for five months when the vaccination coverage had rapidly increased (June to October in 2021), we constructed a large-scale vaccine dataset. Having a fraction of the tweets annotated with the stances of users who posted them on vaccination uptake, we developed a deep neural network on the basis of content and network features to label the user's vaccination stance for each tweet. We next aggregated the predicted stances for each user in the early, middle and late month-long periods to study diachronic changes in the vaccination intention of individuals and ultimately obtained deeper insights into their information-sharing behaviors to determine the factors that affected their stance change.

II. Dataset Construction

Twitter is one of the most popular social media platforms in Japan, with users covering a wide range of age groups, especially young people.[2] Users can submit posts (tweets) they have written as well as react to tweets written by others, called reactions. The reactions, consisting of retweets, quotes, and replies, indicate that the users are interested in the tweets. Thus, we used tweets to identify people's stances towards the vaccination and used the reactions to analyze their information-sharing behaviors.

In Japan, vaccinating people aged 18 or older began on June 17, 2021, and by the end of October, nearly 75% of the population was fully vaccinated. We collected all Japanese tweets[3] from June 1, 2021 to October 31, 2021, and extracted tweets containing the keyword "ワクチン" (wakuchin, vaccine in English). We excluded tweets generated by "share via

* These authors contributed equally.
† Currently, he works for Microsoft, Japan.
[1] https://ourworldindata.org/covid-vaccinations
[2] https://www.humblebunny.com/japans-top-social-media-networks
[3] Twitter data is provided by NTT data.

IEEE/ACM ASONAM 2022, November 10-13, 2022
978-1-6654-5661-6/22/$31.00 © 2022 IEEE

TABLE I: Overview of annotation criteria.

Labels	Criteria
Pro-vaccine	Saying the user was vaccinated Recommending vaccination to others Criticizing anti-vaccine people
Anti-vaccine	Expressing intention not to get vaccinated Calling attention not to get vaccinated Criticizing pro-vaccine people
Neutral	Showing facts (e.g. number of vaccinated) Introducing the press of public institutions Discussing topics irrelevant to pros and cons of vaccination

Twitter" function on some websites and by an application called "shindanmaker." These tweets contain no information to estimate the intent of users. We thus constructed a vaccine tweet dataset with 7,912,014 tweets posted by 1,213,747 unique users.

By investigating the transition in the number of collected tweets, we found that the peak number of tweets occurred on August 26, 2021, when the completion rate of the first vaccination reached 55% of the national population, and the number of tweets decreased as the vaccination progressed.

III. Vaccination Stance Classification

As the basis and key to subsequent analysis, we identified the users' stances towards vaccination. Previous studies on predicting vaccine stance from tweets [11], [12] relied on textual information and failed to classify tweets that referred to posts with the opposite vaccination stance. We additionally used reaction graph information to classify tweets into the users' stances towards vaccination using a deep neural network.

A. Stance annotation of tweets

To train a deep learning model for stance classification, we manually annotated a small subset of vaccine tweets into three class labels: pro-vaccine, anti-vaccine, and neutral to vaccine. To annotate reliable labels, we first developed the criteria of stance annotation, as shown in TABLE I, and measured the annotation-matching rate. In accordance with these annotation criteria, four annotators labeled the same 500 tweets to measure inter-annotator agreement; Fleiss's kappa coefficient [13] on this annotation task was 0.74, which confirms the stability of the annotations. Each annotator then labeled on average 1800 randomly-chosen tweets; thus we obtained a total of 7254 labeled tweets.

B. Text and Graph-based stance classification

With the above annotated tweets, we next trained a deep neural network on the basis of the textual content and reaction to classify vaccine stance. Our model includes three components: text encoder, reaction encoder, and classifier.

The text encoder induces linguistic features from the tweet text. We fine-tuned a pre-trained Bidirectional Encoder Representations from Transformers (BERT) [14] on our target task. For each tweet, we carried out basic preprocessing, such as full-width half-width conversion, case conversion, and removal of various symbols to the input before inputting it to BERT.

A user's stance can be influenced by who that user interacts with. The reaction encoder extracts reactions (retweets, quotes, and replies) between users and generates each user's reaction vector (RA vector) representing who reacted to that user and who the user reacted to. To reduce the computational costs, we used reactions to the most influential users. Specifically, we divided each month into three periods, 1st day to 10th, 11th to 20th, and 21st to 30th (31st), and collected the top-10K users who reacted to others (hereinafter, information spreaders) and the top 10K users who others reacted to (hereinafter, information senders) for each period. We then vectorized the number of reactions between the users and top information spreaders/senders in the last three periods. The obtained RA vector was input to the fully-connected layer and tanh function to reduce the number of dimensions.

The classifier inputs a concatenation of the tweet-text and reaction-graph vectors to a fully-connected layer. It then passes the output to the softmax function to make a prediction.

TABLE II: Vaccine-stance classification: results.

model	macro F_1	F_1 for anti	F_1 for neutral	F_1 for pro
BERT	0.641	**0.457**	0.590	0.878
BERT+RAvec	**0.665**	0.406	**0.695**	**0.893**

C. Experiments on vaccine stance classification

1) Settings: We used 20% of the dataset as the test set and split the rest into training and development sets with the ratio of 8:2. We ensured that the proportion of stance labels in each set was the same; among the 4189 tweets in the training set, there were 2903 pro-vaccine, 186 anti-vaccine, and 1100 neutral tweets.

To implement the text encoder, we used the Japanese BERT pre-learning model released by NICT, Japan.[4] We set the maximum number of tokens to 160. For the reaction encoder, we obtained RA vectors with 500 dimensions by feeding the original 95,016 dimensional vectors to two fully connected layers and the tanh function. To confirm the utility of the reaction information for vaccine-stance classification, we compared the classifier with and without the reaction encoder (BERT and BERT with RA vector).

2) Results: TABLE II lists the results including the macro-F_1 and F_1 scores of each class. BERT+RAvec obtained a higher macro-F_1 score than BERT. This indicates that RA vectors contribute to improving prediction. Because the prediction performance of the anti-vaccine class is not good due to the small number of tweets in the class, we set a probability threshold to obtain reliable labels. When the class with the maximum output of the softmax function was the anti-vaccine class, we increased the threshold to 0.7. Thus, the precision of the anti-vaccine class increased from 0.41 to 0.60, which is

[4] https://alaginrc.nict.go.jp/nict-bert/index.html

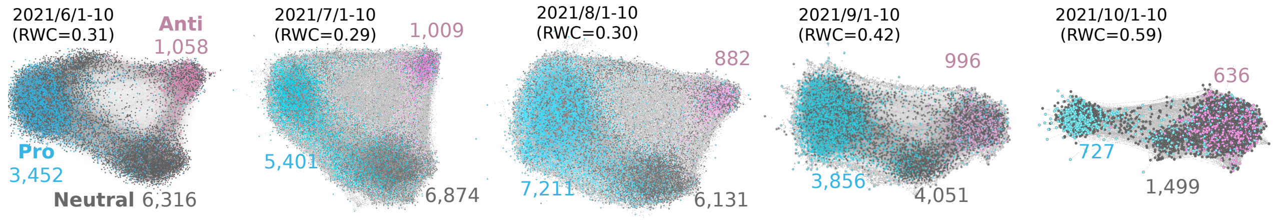

Fig. 1: Evolution of polarization of reaction graphs, RWC, and number of users with each stance.

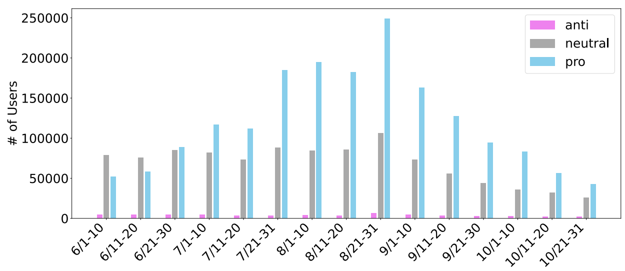

Fig. 2: Changes in number of users with vaccine stances.

not much worse than other classes. Instead, the recall of the class decreased from 0.41 to 0.26.

IV. Analysis

We applied the vaccine-stance classifier to the vaccine-relate tweets within the five months (June to October, 2021) to track the evolution of users' stance towards vaccination and the factors (information the users refer to) that would have affected changes in their vaccine stance.

A. Distribution of users' stance

After all the tweets were labeled using our stance classifier, we aggregated the labels to determine users' stances. We assumed that people do not change their stances in a short period, and divided each month into three periods, to aggregate tweet labels by each user using majority votes. In the case of a tie, the user stance was determined by the priority of pro-, neutral, and anti-vaccine in accordance with the order of prediction precision.

Fig. 2 illustrates the distribution of users' stance in the 15 time periods during the 5 months. The number and proportion of pro-vaccine users increased before the peak number of tweets, and decreased after the second dose. The number of anti-vaccine users was significantly small but consistent.

B. Transition in polarization between stance groups

To determine whether polarization between pro- and anti-vaccine users occurred as reported in online vaccine debates on other infections [11], [15], [16], we visualized the graph of reaction behaviors between users and the distribution of users' stances in the graph as in a previous study [11]. We used an undirected reaction graph in which each node represents a user and each edge represents the existence of reactions between users at both ends. To observe the transition in polarization, we extracted the reaction graphs for the 15 time periods and drew each graph using Gephi.[5] We only depict the nodes with 30 or higher degrees because these nodes represent users who actively shared information.

Fig. 1 illustrates the reaction graphs of the first ten days of each month. The color of nodes indicates the vaccine stance of users. We can see three groups of densely connected nodes in each period, which represent three different stances. Connections between the pro- and anti-vaccine groups are sparse compared with connections between the neutral group and the others. This indicates that the pro- and anti-vaccine groups were consistently polarized through periods.

To quantify the polarization, we used the Random Walk Controversy (RWC) [17] for measuring the degree of separation between two communities in a graph as the ratio of random walks staying in the same community. In each period, we first detected densely connected nodes on the reaction graph as communities using the Louvain method [18].[6] We then assigned either of three stances to the three largest communities by majority vote among stances of users in each community. Finally, we calculated the RWC between pro- and anti-vaccine communities in each period. In this process, we slightly modified the random walk process in the original RWC; starting from a node in either pro- or anti-vaccine community, and stopping at high-degree nodes in either community.

As shown in Fig. 1, the size of the neutral and pro-vaccine group rapidly decreased after the beginning of September, which was when the number of vaccinated people increased rapidly on an unprecedented scale, while the size of the anti-vaccine group was rather consistent. This indicates that most fully vaccinated users did not attend or left the discussions, whereas the anti-vaccine users still continued their activities. The RWC of the reaction graph continued to grow because the number of neutral users, bridging pro-vaccine and anti-vaccine users, significantly decreased.

C. Changes in stance of individual users

We next analyzed changes in the stances of individual users, focusing on active users who were posting repetitively during the data-collection period. Note that the studies on online vaccine debates [6], [11] did not analyze changes in per-user vaccination stances. For all 15 time periods, we extracted 135,784 users who posted once or more in at least 10 periods.

[5] https://gephi.org/

[6] We specify the "resolution" parameter in this method which affects the size of the smallest community to be detected as 2.

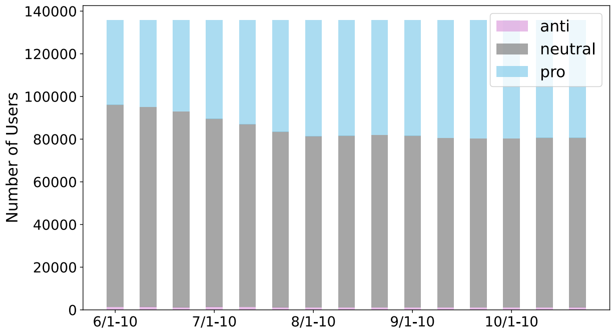

Fig. 3: Changes in user stance distribution.

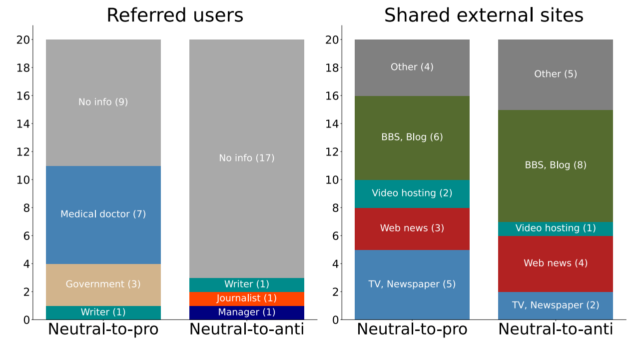

Fig. 4: Users referred to by neutral-to-pro/anti users (left) and external sites shared by neutral-to-pro/anti users (right).

For the period when there was no posting, the stance was labeled the same as the stance of the last period.

Fig. 3 shows the transition of the user stance distribution. The number of anti-vaccine users remained almost unchanged, whereas a certain number of neutral users became pro-vaccine. Among the 92,619 users who were initially neutral at the beginning, 17,488 changed their stance only once, of which 16,734 changed to the pro-vaccine group and 754 changed to the anti-vaccine group. Once the users changed to a pro-vaccine or anti-vaccine group, only a few would switch back to the opposite stance. These results suggest that people who were initially hesitant to get vaccinated ended up getting vaccinated.

D. Factors behind change in user stance

As shown above, since there was almost no transition between the pro- and anti-vaccine groups, we focused on the users whose stance was initially neutral and changed once to pro-vaccine (neutral-to-pro) or anti-vaccine (neutral-to-anti). We analyzed their information-sharing behaviors from the view point of users and external sites.

1) Which users are referred?: We investigated which user accounts were typically referred to (replied, retweeted, or quoted) by the neutral-to-pro and neutral-to-anti users. We first collected user accounts referred to by these users for 20 days including the period that they changed stance and the last period before the change. We also collected user accounts referred by 36,947 neutral users who stuck to their stances (remaining-neutral). To extract user accounts typically referred to by the neutral-to-pro (or neutral-to-anti) users, we used the chi-squared test of independence on two user account groups referred to by the neutral-to-pro (or neutral-to-anti) users and the remaining-neutral users at a significance level of 1%. After finding the user accounts typically referred to by each group, we further investigated the attributes of these accounts such as occupation.

The left of Fig. 4 shows the attributes of the top-20 user accounts referred to by the neutral-to-pro and -anti users. Comparing the two graphs, the user accounts referred to by neutral-to-pro users included 7 accounts of medical doctors and 2 accounts of governments, while such users are rarely seen in the neutral-to-anti user group. This indicates that users often refer to announcements and comments from medical doctors and governments when deciding to get vaccinated.

2) Which types of external sites are shared?: We next investigated the external sites that were shared by neutral users. We collected the URLs and headlines of the tweets containing the external sites for neutral-to-pro, -anti and remaining-neutral users. If one user shared the same link multiple times, it was counted only once. Similar to the referred user account analysis, we investigated typically shared external sites by the neutral-to-pro/anti users by conducting a chi-squared test of independence.

The types of the top-20 sites shared by the neutral-to-pro and -anti users are shown on the right of Fig. 4. Among the sites shared by neutral-to-pro users, eights (40%) are mass and web media including TV, newspaper, and web news sites. For neutral-to-anti users, we found they shared BBS and blog sites more than neutral-to-pro users.

Fig. 5 illustrates the word cloud using the headlines of the shared links. For neutral-to-pro users, we found the word "reservation" from July because they checked and shared the information on how to make a vaccination reservation or tweeted that they had made a reservation. In September, the word "Taro Kono" refers to the minister in charge of vaccination at that time, who posted an article to criticize the misinformation about vaccines on his official website. Taken together, we found that neutral-to-pro users referred to the latest news and opinions of public figures.

The word cloud of the neutral-to-anti users always contained the word "cause," which is accompanied by the word "death" or "damage," suggesting that the neutral-to-anti users were particularly concerned about the vaccine's safety. In July, the words "immunity" and "destruction" appeared, which is considered to be from the view that booster shots every four months could eventually weaken the immune response. Another feature in the word cloud of neutral-to-anti users is the appearance of words reminiscent of public institutions and experts, such as "World Health Organization (WHO)" and "Ministry of Health, Labour and Welfare (MHLW)." However, after further investigation of the content, we suspect that the blogs that over-interpret expert opinion may have influenced the neutral-to-anti users. In short, while the neutral-to-pro

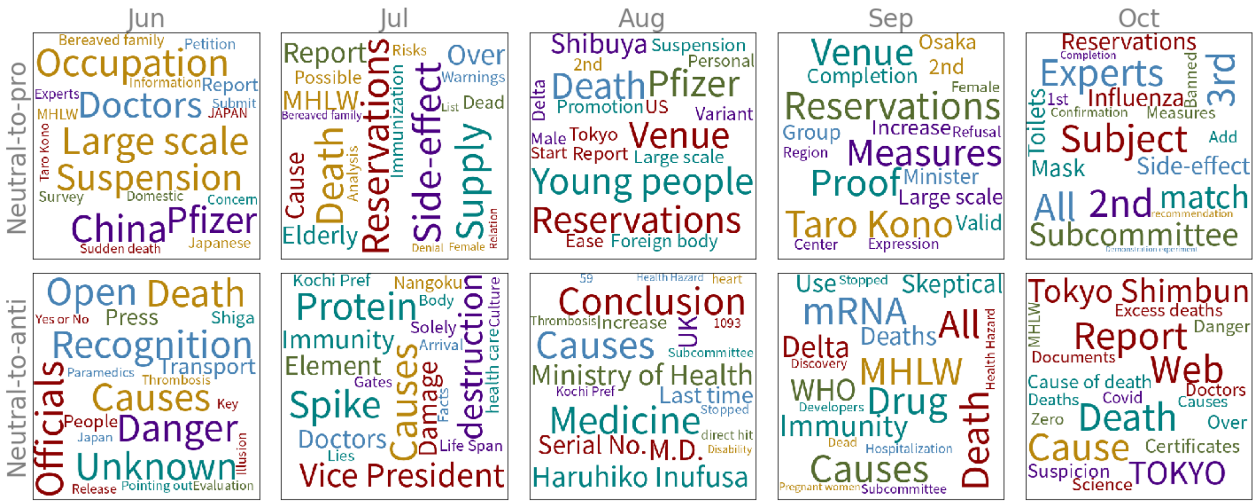

Fig. 5: Changes in titles of external sites referred to by neutral-to-pro users (top) and neutral-to-anti users (bottom).

users referred to the latest information from news and public institutions, the neutral-to-anti users mainly took up the reports without sufficient discussion.

V. Conclusions

To learn from the successful COVID-19 vaccination campaign in Japan, we analyzed diachronic changes in stances of Twitter users towards COVID-19 vaccination. We developed a BERT-based stance classifier with reaction information and applied it to all vaccine-related tweets from June to October, 2021. We found that users' stances were polarized and neutral users had much reactions with pro-vaccine users. Thus neutral users who changed their stance turned pro-vaccine in most cases. This indicates that guiding neutral users to get vaccinated is key to improving the vaccination coverage. When the users turned pro-vaccine, they referred the account of doctors and governments or shared news from mass media, suggesting the importance of accurate and credible publicity.

Acknowledgments

This research was conducted as part of "COVID-19 AI & Simulation Project" run by Mitsubishi Research Institute commissioned by Cabinet Secretariat, JAPAN. The methods for analysis were developed with support from JST CREST Grant Number JPMJCR19A4 and JSPS KAKENHI Grant Number JP21H03445.

References

[1] M. Sallam, "COVID-19 vaccine hesitancy worldwide: a concise systematic review of vaccine acceptance rates," *Vaccines*, vol. 9, no. 2, 2021.
[2] A. De Figueiredo, C. Simas, E. Karafillakis, P. Paterson, and H. J. Larson, "Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study," *The Lancet*, vol. 396, no. 10255, pp. 898–908, 2020.
[3] S. Nomura, A. Eguchi, D. Yoneoka, T. Kawashima, Y. Tanoue, M. Murakami, H. Sakamoto, K. Maruyama-Sakurai, S. Gilmour, S. Shi, H. Kunishima, S. Kaneko, M. Adachi, K. Shimada, Y. Yamamoto, and H. Miyata, "Reasons for being unsure or unwilling regarding intention to take COVID-19 vaccine among Japanese people: A large cross-sectional national survey," *Lancet Reg. Health West. Pac.*, vol. 14, 2021.
[4] E. Dubé, M. Vivion, and N. E. MacDonald, "Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications," *Expert Rev. of Vaccines*, vol. 14, no. 1, pp. 99–117, 2015.
[5] M. Thelwall, K. Kousha, and S. Thelwall, "COVID-19 vaccine hesitancy on English-language Twitter," *Prof. Inf.*, vol. 30, no. 2, 2021.
[6] S. Liu, J. Liu *et al.*, "Understanding behavioral intentions toward COVID-19 vaccines: theory-based content analysis of tweets," *J. Med. Internet Res.*, vol. 23, no. 5, 2021.
[7] M. T. J. Ansari and N. A. Khan, "Worldwide COVID-19 vaccines sentiment analysis through Twitter content." *Electron. J. Gen. Med.*, vol. 18, no. 6, 2021.
[8] F. Cascini, A. Pantovic, Y. A. Al-Ajlouni, G. Failla, V. Puleo, A. Melnyk, A. Lontano, and W. Ricciardi, "Social media and attitudes towards a COVID-19 vaccination: A systematic review of the literature," *EClinicalMedicine*, vol. 48, 2022.
[9] T. Hu, S. Wang, W. Luo, M. Zhang, X. Huang, Y. Yan, R. Liu, K. Ly, V. Kacker, B. She *et al.*, "Revealing public opinion towards COVID-19 vaccines with Twitter data in the united states: spatiotemporal perspective," *J. Med. Internet Res.*, vol. 23, no. 9, 2021.
[10] K. Miyazaki, T. Uchiba, F. Toriumi, K. Tanaka, and T. Sakaki, "Retrospective analysis of controversial topics on COVID-19 in Japan," in *Proc. ASONAM*, 2021, pp. 510–517.
[11] X. Yuan, R. J. Schuchard, and A. T. Crooks, "Examining emergent communities and social bots within the polarized online vaccination debate in Twitter," *Soc. Media Soc.*, vol. 5, no. 3, 2019.
[12] L.-A. Cotfas, C. Delcea, I. Roxin, C. Ioanăş, D. S. Gherai, and F. Tajariol, "The longest month: analyzing COVID-19 vaccination opinions dynamics from tweets in the month following the first vaccine announcement," *IEEE Access*, vol. 9, pp. 33 203–33 223, 2021.
[13] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.
[14] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.
[15] A. L. Schmidt, F. Zollo, A. Scala, C. Betsch, and W. Quattrociocchi, "Polarization of the vaccination debate on Facebook," *Vaccine*, vol. 36, no. 25, pp. 3606–3612, 2018.
[16] A. Cossard, G. D. F. Morales, K. Kalimeri, Y. Mejova, D. Paolotti, and M. Starnini, "Falling into the echo chamber: the Italian vaccination debate on Twitter," in *Proc. ICWSM*, 2020, pp. 130–140.
[17] K. Garimella, G. D. F. Morales, A. Gionis, and M. Mathioudakis, "Quantifying controversy on social media," *ACM Trans. Social Comput.*, vol. 1, no. 1, pp. 1–27, 2018.
[18] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech: Theory Exp.*, vol. 2008, no. 10, 2008.