

Depicting U.S.-China Disputes on Tech Giants through Social Media: An Attempt of Computational Political Communication

Anonymous Submission

Abstract— Social media has become a promising field for data scientists to carry out collaborative and interdisciplinary researches with social scientists. Computational political communication based on big data analytics of social media texts brings a prospecting framework to understand the public's perception of and interaction with certain political issues on a global basis. This study reviews previous efforts and offers a preliminary analysis of our collected Twitter data, delineating online political communication dynamics upon U.S.-China disputes, a most influential global agenda today. Tech giants Huawei, Tencent and ByteDance are chosen as epitomes of U.S.-China power game to grasp more detailed opinions. Seven English-speaking countries: U.S., U.K., Canada, Australia, New Zealand, India, and Pakistan, along with the three companies, are selected as keywords for filtering among all the Twitter data collected from March 2020 to September 2020 across the globe. Our demo shows that the popularity of discussions about a particular country and company is not consistent, and might be event-induced. Also, the discourse of all these companies are inter-related rather than separated. This research facilitates future study, including fine-grained and categorized sentiment analysis and the generalization of VICs to depict online public opinions.

Keywords— Big data, Social media, Computational political communication, U.S.-China disputes, Tech giants, Twitter

I. INTRODUCTION

Big data analytics and social media are shedding new light on the interdisciplinary study of political communication. For the first time in history, researchers are faced with endless political information and comments generated by ordinary citizens on social media, empowering them to explore the dynamics of public opinion on critical issues.

U.S.-China bilateral relation has become one of the most influential agenda in contemporary global politics. Over the past few years, started from increasing mutual tariffs, to the escalating conflicts embodied in almost every aspect of the two countries' interaction including international governance and national security, U.S.-China disputes have drastically altered the ecosystem for global trade, investment and supply chains. Most specifically, from the blacklisting of semi-conductor suppliers for Huawei, to the ban of TikTok alongside with WeChat, tech giants with a Chinese background are to a great extent posed as epitomes of the competitive power game scenario [1-2].

Despite its friendly and goodwill showbiz in the first two years, the Trump administration has then embedded a pivot of America's policy towards China. Published in December 2017, "National Security Strategy of the United States of America" initiates Washington's novel definition of China as a "competitor" "rival" "adversary" and "revisionist power", which served as curtain-raiser for the upheaval until this day [3]. With rising nationalism domestically and intensifying pressure externally, Chinese decision makers, on the other hand, have also been adopting a more "assertive" and "uncompromising" approach in both *Aussenpolitik* and *Innenpolitik* which inevitably account for de facto "lex talionis" deadlock [4].

The post-COVID-19 pandemic era has witnessed an exacerbated "decoupling" between the two major powers [5]. A vis-à-vis confrontation and retaliation has been projected to not only the sphere of trade but more structurally impacting the network of "global tech" constructed via cooperative efforts within last two decades. As a milestone of on-going industrial revolution, new phase of globalization and forefront of free trade, international tech giants have their roots in crossing-border information exchange, and simultaneously contribute to deepening the connection of "global village" [6-8].

Huawei, with its headquarter located in Shenzhen, is a representative telecommunication hardware manufacturer and flagship smart device company of China. In 2020, this 5G titan has ultimately met with the suppression from U.S. featuring the forced outage of its semi-conductor suppliers including TSMC, Samsung and SK Hynix [9].

China-based software companies have also been dramatically impacted in 2020 by the "Clean Network" action plan of U.S. government. TikTok, a social media app for short videos of ByteDance, and WeChat, another app for instant message, audio and video calls owned by Tencent, are recently included in the blacklist of Washington. Not until these latecomer high-tech innovators enjoyed their unprecedented international boom, they are confronted with the strike that may end their game in the U.S. or even all U.S. allies [10].

This study looks into Huawei, ByteDance and Tencent for their representativeness of hardware and software providers impacted by U.S.-China disputes, and for their huge user scale and global influence. Moreover, WeChat has its major income from Chinese Mainland while TikTok is an international market-oriented app. This can also serve as a comparison.

Internet tech giants are considered to be the watershed of China's rise as a global competitor in technology, innovation and tertiary industry whereas at the same time an epicenter of China's "challenge" and "threat" to the U.S..

In this digital era, while state behaviors, state-company relations and geopolitical power plays can be more precisely and directly depicted from macroscopical policy-making level, the public's perception and interaction also constitute a fundamental sector to the decision-making process, possibly more profoundly than ever. Twitter, among all major global social media platforms, has its characteristics as an internationally connected, weak ties-based "public sphere", facilitates this study to depict a more comprehensive view of U.S.-China disputes [11]. Therefore, this study chooses Twitter data for delineating online political communication dynamics.

In this study, names of seven major English speaking countries, namely, the United States, the United Kingdom, Canada, Australia, New Zealand, India, and Pakistan, are selected as keywords for filtering among all the Twitter data collected over six months across the globe. The U.S., UK, Canada, Australia and New Zealand are also known as "Five Eyes Alliance", which is an intelligence-based strategic ally group with its origin as "UKUSA" from the Second World War.[12] This study also notices the strong geopolitical proximity between India, Pakistan and China. These two South Asian countries, acknowledged as the biggest English speaking countries adjacent to China and together populated over 1.5 billion, cast huge impact on China's overseas market and foreign strategies. [13] Names of the three Chinese tech giants, Huawei, Tencent, and ByteDance, are also used to target relevant tweets.

How does online public opinion perceive and respond to the disputes between America and China? What is the panorama of social media discourse on specific tech giants? To our knowledge, this is the first study dedicated to answering these questions. In the second section, we will provide a concise review of the transition of political communication studies, and point out its promising future for both social and data scientists. Then, we go on to introduce preliminary findings of our Twitter data. Discussions and conclusions are offered at the end. We hope that our publication of the dataset and its primitive findings can arouse more scholars to participate in further research.

II. EXISTING WORKS: ACHIEVEMENTS, DEFICIENCIES, AND EXPECTATIONS

This section summarizes existing works in understanding the dynamics of public opinion, pointing out its shift from traditional media to social media and the impetus behind it. We show that large-scale social media data ensures a promising future for the interdisciplinary study of political communication as well as social simulating and modelling.

A. *Traditional media and surveys in the beginning: Efforts by social scientists*

Media framing of certain countries, issues or important figures have long been one of the most popular topics among political communication researchers. The manually-coded

methods, e.g., text analysis, discourse analysis and content analysis, are widely adopted in the studies of images and agenda constructed by different media sources. While a majority of the researches by social scientists still focus on traditional media, namely newspapers, magazines, TVs, etc.

In recent years, China has won more and more attention to its international image presentation. Golan and Lukito's [14] describes the rise of China via looking into American Newspaper' "opinion articles" on China. Golan et al. reckons that opinion journalism, including "editorials" and "co-eds" of major newspapers, embodies the most clearly how a country's elites community is thinking about certain issues. This research mainly utilizes the inductive qualitative framework of analysis, looking into the Wall Street Journal (WSJ) editorials and the op-eds on China's rise. The text analysis is mainly about key statements, positive or negative attitudes and use of quotes. It turns out that "economic partnership", "internal dispute", "geopolitical threat" and "economic threat" are the main categories of WSJ's narrative framework. Golan et al. carried out the analysis based on a belief that the elites of society are responsible and influential to foreign policymaking. Similar frameworks are universally adopted in social scientist's research on media framing and international image-related topics.

While public opinion's response to certain countries and issues are becoming more well-attended in academia these days, classic methodology including surveys are applied in numerous social science researches. Yang [15] carried out the study on how China's image affect the selling of Chinese product in the United States. Focusing on public opinion as a key factor in this research, participants were recruited in a medium city of Ohio in both ways: printed and online questionnaires. However, this approach to depict public opinion about China's country-of-origin still faces with questions that, is it comprehensive enough, and is it the most efficient way to do the research?

The Verb In Context System (VICS) is another early attempt to understand the mentality of people, especially political leaders [16]. Based on the 1969 study of the "operational code" [17], the VICS analyses people's political opinion and beliefs on power, predictability, role of chance, etc., through their use of verbs found in public accessible speeches and policy text. An exhaustive dictionary was established to provide reference to the orientation of each verb, e.g., friendly or hostile, optimistic or pessimistic. The VICS is widely used for study political figures [18-22], while its generalization to the public remains somehow stagnant, at least partly due to a lack of available texts written by ordinary citizens.

B. *The advent of social media: Promises for data scientists*

Entering the era of social media, the openness and user-generating nature of cyberspace empowers general public to express and construct online discourse. According to a Pew Research Center report, more than 40% of American adults accessed to information for the 2016 presidential election via social media, which provides a simple example for this ubiquitous phenomena nowadays [11]. The rapid development

of computer science enables data scientists to directly study the public's opinion for the first time. Opposite to social scientists' focus on limited information sources and target population, data science researchers has been working to discover more latent and complex information from broader social media data.

Sentiment analysis is a frequently employed technique for studying social media data in text modality. The basic goals of sentiment analysis are emotion recognition and polarity detection [23, 24]. Many research used this method to explore country images, evaluate international relations, and predict electoral results.

Xu et al. [25] and Chen et al. [26] are both event-based country image study with Twitter data, observing online public opinion during the 70th anniversary of the People's Republic of China and the COVID-19 pandemic, respectively. Their data was retrieved through Twitter Streaming API, and sentiments towards China (positive, negative, neutral) were analysed with machine learning algorithms trained on manually labelled data. Their features include: 1) Xu et al. collected and compared English and Chinese data, while Chen et al. focused on English discourse. It was found that a significant opposition existed between the online public opinions towards China of the two languages. 2) Chen et al. provided fine-grained sentiment analysis by dividing online public opinion towards China into seven categories: Politics, Economy, Foreign affairs, Culture, Epidemic situation, Anti-epidemic measures, and Racism. They revealed that the gradual increase in negative politics-, foreign affairs-, and racism-related tweets and the decrease in non-negative epidemic situation-, anti-epidemic measures-related tweets resulted in the overall sentiments' transition from non-negative to negative towards China. 3) Chen et al. displayed the different patterns in the attitudes of Congress members, media, and social bots, showing that social bots were more likely to spread negative sentiments towards China, while media were usually non-negative. For U.S. congress members, the Republicans were more negative than the Democrats. 4) Xu et al. explored how positive and negative tweets were distributed among different countries and found that states enjoying better diplomatic relations with China generally had a positive view towards China. 5) Xu et al. obtained word vectors for the top 100 frequently and uniquely used words for both English and Chinese, positive and negative tweets through word2vec. Preferred topics of distinct languages and sentiments were analysed, e.g., positive Chinese tweets mostly focused on celebration activities while negative Chinese tweets tended to talk about broader topics like Hong Kong.

Chambers et al. [27] modelled relations between nation states using sentiments revealed in tweets with country names. Seventeen months of Twitter data was collected and aggregated sentiments for nation pairs were calculated with support vector machine. The results indicated an alignment between Twitter sentiments and human polls & official diplomatic relations, verifying the validity of applying social media data to infer international relations.

Predicting election results with social media data is also a focus for researchers. Related works include [28-31]. Other

TABLE I. TABLE TYPE STYLES

Type	Names and synonyms
Countries	'the us', 'UnitedStates', 'United States', 'the states', 'America', 'uk', 'UnitedKingdom', 'United Kingdom', 'Britain', 'Canada', 'Australia', 'aussie', 'NewZealand', 'New Zealand', 'India', 'Pakistan'
Companies	'huawei', 'hua wei', 'bytedance', 'byte dance', 'zijietiaodong', 'zi jie tiao dong', 'tiktok', 'tik tok', 'douyin', 'dou yin', 'tencent', 'tengxun', 'teng xun', 'wechat', 'weixin', 'wei xin'

papers addressing online public opinion towards political events include [32-34].

C. The era of interdisciplinary collaboration: Computational political communication

Despite the considerable endeavour and contributions the above-mentioned works made to the emerging field of computational political communication, the shortcomings of them are also prominent. Their implications to social challenges are vague. With their vision limited to describing general pictures, the advanced computational techniques are not fully utilized to answer more meaningful questions and bring about possible solutions. Meanwhile, a lack of real-time, or 'nowcast', analysis, which has the potential to detect major events at an early stage and provides governments and the society with necessary notifications, also stands out as a major deficiency for existing studies.

From the global communication perspective, it can be anticipated that a more comprehensive and real-time computational research on social media will become increasingly significant for academia and policy-makers. Computational political communication is undoubtedly a rising field for interdisciplinary collaboration, with social scientists' intrinsic dedication to find questions and create meanings and data scientists' capability to initiate more sophisticated quantitative researches. Data scientists should be encouraged to engage in more of these interdisciplinary research, honing and experimenting their methodologies and theories [35, 36].

III. DATA COLLECTION AND ANALYSIS

This section introduces a Twitter dataset and provides a demo for analysing political information on social media.

A. Data Collection

The data used in this study is collected through Twitter Streaming API, which allows users to retrieve tweets with designated hashtags. Since the dataset is expected to support research in a broader context and is not specific to this study, the hashtags for retrieving data are designed to include all major countries in the world (5 permanent members in the UN security council, G20 countries, OECD countries) and countries enjoying close contact with China (member states of Shanghai Cooperation Organization and the Association of Southeast Asian Nations, North Korea). The collection started from March 4, and will continue until next march. This study uses the data from March 4 to September 14, 2020.

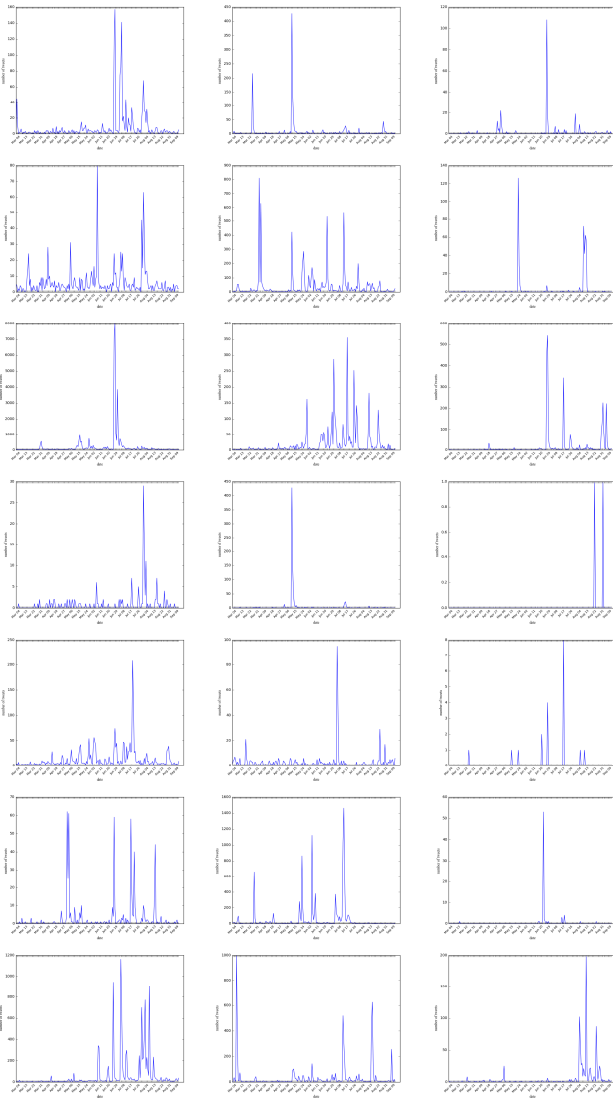


Fig. 1. Number of tweets discussing the seven countries and the three companies, from March 4 to September 14. (from top to down: Australia, Canada, India, New Zealand, Pakistan, the United Kingdom, the United States; from left to right: ByteDance, Huawei, Tencent)

Later, we filter the collected data with the following criteria: a tweet should simultaneously include names or synonyms of at least one of the seven English speaking countries and the names or synonyms of at least one of the three Chinese tech giants in order to be kept. The names and synonyms are shown in Table I. Please note that the names and synonyms are case-insensitive, since all the tweets and names would be changed into lower case before filtering. At last, a total of 83,597 tweets were selected for this study.

B. Data Analysis

As a preliminary study, we will reveal quantitative characteristics and linguistic preferences of tweets linked to different countries and companies.

Fig. 1 displays the number of tweets talking about the seven chosen countries and the three selected companies during March 4 and September 14. It can be found that the ups and

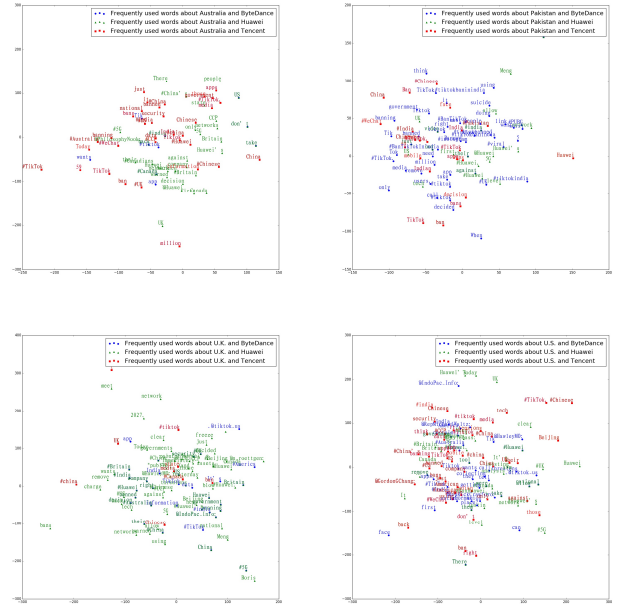


Fig. 2. Word vectors of frequently mentioned words. (from top to down and left to right: Australia, Pakistan, U.K., U.S.)

downs in popularity of country-company topics are significant: rarely mentioned in most days and intensively discussed in certain periods, possibly after an important event. Some events may induce massive discussions related to the same company in multiple countries, e.g., following India's ban on TikTok on 29 June, a peak of related tweets appear in Australia, India, U.K., and U.S.. Also, public attention towards the three companies are uneven. Huawei receives the most mentions when people are simultaneously talking about Australia, Canada, New Zealand, and U.K., while ByteDance enjoys the popularity in India, Pakistan, and U.S.

Fig. 2 reveals the word vectors of frequently mentioned words (appeared more than 1,000 times in our collected corpus) by tweets related to different countries and companies. Limited by space, this paper only shows the word vectors of Australia, Pakistan, U.K., and U.S.-related tweets. The word vectors were calculated through word2vec. It can be observed that, for any given country, words from tweets related to different companies generally form only one cluster, entangling with each other, rather than distribute separately. It indicates that discourse of all these companies are inter-related. The colors also provide an intuitive image of which company is more widely discussed, e.g., the prominent green in the U.K. picture tells that Huawei was overwhelmingly talked about.

IV. CONCLUSIONS AND DISCUSSIONS

The bilateral relation between U.S.-China in the past few years have been witnessing a "freefall", posing concern to global governance and international order. This study highlights the prospect computational political communication has for people's understanding of such major political events in the Internet era. Social media, with its crossing-border and user-generated nature, provides a promising future for understanding the dynamics of online public opinion about global politics, and the interaction between them.

After summarizing researchers' transition from traditional media and survey to large-scale social media data, we introduce a new Twitter dataset and provide a primitive demo for the use of such data. Future research will include fine-grained sentiment analysis to discover the fluctuation and stabilization of the political leanings of the public. Also, generalizing the VICS to public discourse worth in-depth study.

REFERENCES

- [1] M. Goodman et al., "Perspectives on the Global Economic Order in 2019: A U.S.-China Essay Collection," Center for Strategic and International Studies (CSIS), 2019.
- [2] C. Rosset, "Huawei Ban Means the End of Global Tech," *Foreign Policy*, vol. 17, 2019.
- [3] "National Security Strategy of the United States of America," The White House, Washington, 2017.
- [4] M. Clarke, "Beijing's Pivot West: The Convergence of Innenpolitik and Aussenpolitik on China's 'Belt and Road'," *Journal of Contemporary China*, vol. 29, no. 123, pp. 336-353, 2020.
- [5] T. Bernes et al., "Challenges of Global Governance Amid the COVID-19 Pandemic," *Council on Foreign Relations*, 2020.
- [6] A. Fatma and N. Bharti, "Perception vs. reality: understanding the US-China trade war," *Transnational Corporations Review*, vol. 11, no. 4, pp. 270-278, 2019.
- [7] A. Wyne, "How to Think about Potentially Decoupling from China," *The Washington Quarterly*, vol. 43, no. 1, pp. 41-64, 2020.
- [8] J. Bareviciute, "The locality of the 'global village' in the aspect of communication: Pro et contra m. McLuhan," *Limes: Cultural Regionalistics*, vol. 3, no. 2, pp. 184-194, 2010.
- [9] L. Guo, M. Y. Zhang, M. Dodgson, and D. Gann, "Huawei's catch-up in the global telecommunication industry: innovation capability and transition to leadership," *Technology Analysis & Strategic Management*, vol. 31, no. 12, pp. 1395-1411, 2019.
- [10] "The Clean Network." Secretary of State. <https://www.state.gov/the-clean-network/> (accessed Sep 21, 2020).
- [11] C. B. Williams, "Introduction: Social media, political marketing and the 2016 US election," ed: Taylor & Francis, 2017.
- [12] S. Albers, F. Wohlgezogen, and E. J. Zajac, "Strategic alliance structures: An organization design perspective," *Journal of Management*, vol. 42, no. 3, pp. 582-614, 2016.
- [13] Y. Sun, "China and South Asia Crisis Management in the Era of Great Power Competition," Norwegian Institute for International Affairs (NUI), 2020. Accessed: 21 Sep. 2020. [Online]. Available: www.jstor.org/stable/resrep25750
- [14] G. J. Golan and J. Lukito, "The rise of the dragon? Framing China's global leadership in elite American newspapers," *International Communication Gazette*, vol. 77, no. 8, pp. 754-772, 2015.
- [15] C. Yang, "How China's image affects Chinese products in a partisan-motivated US market," *Global Media and China*, p. 2059436420922702, 2020.
- [16] M. Schafer and S. G. Walker, "Operational code analysis at a distance: The verbs in context system of content analysis," in *Beliefs and leadership in world politics*: Springer, 2006, pp. 25-51.
- [17] A. L. George, "The 'operational code': A neglected approach to the study of political leaders and decision-making," *International studies quarterly*, vol. 13, no. 2, pp. 190-222, 1969.
- [18] S. G. Walker, "Anticipating attacks from the operational codes of terrorist groups," *Dynamics of Asymmetric Conflict*, vol. 4, no. 2, pp. 135-143, 2011.
- [19] J. Renshon, "When public statements reveal private beliefs: Assessing operational codes at a distance," *Political Psychology*, vol. 30, no. 4, pp. 649-661, 2009.
- [20] J. Renshon, "Stability and change in belief systems: The operational code of George W. Bush," *Journal of Conflict Resolution*, vol. 52, no. 6, pp. 820-849, 2008.
- [21] S. B. Dyson, "Alliances, domestic politics, and leader psychology: Why did Britain stay out of Vietnam and go into Iraq?," *Political Psychology*, vol. 28, no. 6, pp. 647-666, 2007.
- [22] E. Cuhadar, J. Kaarbo, B. Kesgin, and B. Ozkececi - Taner, "Personality or role? Comparisons of Turkish leaders across different institutional positions," *Political Psychology*, vol. 38, no. 1, pp. 39-54, 2017.
- [23] E. Cambria, "Affective Computing and Sentiment Analysis," *Ieee Intelligent Systems*, vol. 31, no. 2, pp. 102-107, Mar-Apr 2016, doi: 10.1109/mis.2016.31.
- [24] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Information Fusion*, vol. 37, pp. 98-125, 2017, doi: 10.1016/j.inffus.2017.02.003.
- [25] Y. Xu, Q. He, and S. Ni, "Understanding Online Public Sentiments: A Machine Learning-Based Analysis of English and Chinese Twitter Discourse during the 2019 Chinese National Day," Manuscript accepted by The 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC 2020).
- [26] H. Chen et al., "Country Image in COVID-19 Pandemic: A Case Study of China," *IEEE Transactions on Big Data*, pp. 1-1, 2020, doi: 10.1109/TBDATA.2020.3023459.
- [27] N. Chambers et al., "Identifying political sentiment between nation states with social media," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 2015, pp. 65-75.
- [28] N. Tsirakis, V. Pouloupoulos, P. Tsantilas, and I. Varlamis, "Large scale opinion mining for social, news and blog data," *Journal of Systems and Software*, vol. 127, pp. 237-248, 2017.
- [29] A. Jungherr, H. Schoen, O. Posegga, and P. Jürgens, "Digital trace data in the study of public opinion: An indicator of attention toward politics rather than political support," *Social Science Computer Review*, vol. 35, no. 3, pp. 336-356, 2017.
- [30] J. C. A. D. Lopez, S. Collignon-Delmar, K. Benoit, and A. Matsuo, "Predicting the brexit vote by tracking and classifying public opinion using twitter data," *Statistics, Politics and Policy*, vol. 8, no. 1, pp. 85-104, 2017.
- [31] E. D'Andrea, P. Ducange, A. Bechini, A. Renda, and F. Marcelloni, "Monitoring the public opinion about the vaccination topic from tweets analysis," *Expert Systems with Applications*, vol. 116, pp. 209-226, 2019.
- [32] S. C. McGregor, "Social media as public opinion: How journalists use social media to represent public opinion," *Journalism*, vol. 20, no. 8, pp. 1070-1086, 2019.
- [33] N. J. Adams-Cohen, "Policy Change and Public Opinion: Measuring Shifting Political Sentiment With Social Media Data," *American Politics Research*, vol. 48, no. 5, pp. 612-621, 2020.
- [34] A. D. Leong and S. S. Ho, "Perceiving online public opinion: The impact of Facebook opinion cues, opinion climate congruency, and source credibility on speaking out," *New Media & Society*, p. 1461444820931054, 2020.
- [35] D. B. Margolin, "Computational contributions: a symbiotic approach to integrating big, observational data studies into the communication field," *Communication Methods and Measures*, vol. 13, no. 4, pp. 229-247, 2019.
- [36] W. van Atteveldt and T.-Q. Peng, "When communication meets computation: Opportunities, challenges, and pitfalls in computational communication science," *Communication Methods and Measures*, vol. 12, no. 2-3, pp. 81-92, 2018.