

PENGUMPULAN DATA TWITTER TENTANG COVID-19 DI INDONESIA UNTUK MENGHITUNG TINGKAT ENGAGEMENT PENGGUNA

Luvia Friska Narulita^{*1}, Dwi Harini Sulistyawati²

^{1,2}Universitas 17 Agustus 1945 Surabaya
Email: ¹luvia@untag-sby.ac.id, ²dwiharini@untag-sby.ac.id
^{*}Penulis Korespondensi

(Naskah masuk: 19 Januari 2021, diterima untuk diterbitkan: 09 Juni 2021)

Abstrak

Dari Twitter, pengguna dapat mendapatkan informasi tentang berbagai hal, begitu juga informasi tentang COVID-19. Dalam penelitian ini, dilakukan penelitian tentang tingkat *engagement* pengguna terhadap tweet yang membahas tentang COVID-19. Penelitian dilakukan terhadap data tweet pada bulan April 2020 hingga bulan Juni 2020. Data tweet didapatkan dari akun-akun yang khusus membahas tentang COVID serta akun milik pemerintah dengan melakukan penyaringan pada tweet yang khusus membahas tentang COVID saja. Tujuan dari penelitian ini adalah untuk mengetahui tingkat *engagement* pengguna yang dapat diartikan dengan tingkat interaksi pengguna terhadap tweet yang membahas tentang COVID-19 pada saat awal virus COVID-19 merebak hingga pada saat mulai dikenalkan istilah *New Normal* atau adaptasi kebiasaan baru. Dari penelitian yang dilakukan, didapatkan hasil bahwa tingkat interaksi pengguna terhadap tweet tentang COVID menurun pada bulan Juni ketika kasus COVID semakin bertambah dan pelaksanaan kebijakan bekerja dari rumah sudah dilaksanakan sekitar dua bulan. Penurunan tingkat *engagement* pengguna sebesar 0,5% pada bulan Juni.

Kata kunci: *twitter, engagement, covid-19*

TWITTER DATA COLLECTION ABOUT COVID-19 IN INDONESIA TO CALCULATE USER ENGAGEMENT LEVELS

Abstract

Twitter users in Indonesia quite a lot. From Twitter, users can get information about various things, as well as information about COVID-19. In this study, research was conducted on the level of user engagement on tweets that discussed COVID-19. The study was conducted on tweet data from April 2020 to June 2020. The tweet data was obtained from accounts specifically discussing COVID and government-owned accounts by filtering on tweets which specifically discussed COVID only. The purpose of this study is to determine the level of user engagement which can be interpreted by the level of user interaction with tweets that discuss COVID-19 at the beginning of the COVID-19 virus spread until it was introduced to the term *New Normal*. This term is introduced to people to encourage them wearing mask and implementing health protocol during their activities. From the research conducted, it was found that the level of user interaction with tweets about COVID decreased in June when COVID cases increased and the implementation of the work from home policy had been carried out for about two months.

Keywords: *twitter, engagement, covid-19*

1. PENDAHULUAN

Pandemi COVID-19 yang melanda banyak negara di dunia, telah menumbuhkan beberapa akun twitter baru yang mengkhususkan diri untuk memberikan informasi terkait pandemi kepada pengguna. Selain akun twitter khusus, akun twitter resmi pemerintah juga hadir untuk memberikan informasi terkini terkait pandemi yang terjadi. Informasi mengenai

jumlah penderita, jumlah pasien meninggal ataupun sembuh, informasi tentang rumah sakit rujukan maupun informasi tentang pencegahan dan kampanye hidup bersih banyak disebarkan melalui akun-akun Twitter tersebut.

Penggunaan Twitter tidak lepas dari adanya perkembangan teknologi dan internet yang memungkinkan orang untuk berkomunikasi melalui sosial media. Penggunaan sosial media memang cukup tinggi di Indonesia, berdasarkan riset yang

dilakukan oleh Wearesocial Hootsuite yang dirilis pada Januari 2019 dan ditulis ulang oleh katadata.co.id (katadata,2020), pengguna media sosial di Indonesia mencapai 150 juta jiwa. Pada riset yang dikeluarkan pada awal tahun 2020, pengguna twitter mencapai 56% dari keseluruhan pengguna media sosial di Indonesia. Tingginya pengguna media sosial di Indonesia, tentunya meningkatkan potensi ekonomi digital di tanah air. Selain itu, penggunaan media sosial untuk mendapatkan informasi telah banyak dilakukan oleh masyarakat di Indonesia. Berdasarkan riset yang dilakukan oleh Chandra Marsono pada tahun 2014 dengan melakukan gabungan riset kualitatif dan kuantitatif dengan memanfaatkan platform online yaitu Survey Monkey dan Google Form, didapatkan hasil bahwa pengguna internet di Indonesia pada kategori dewasa muda mulai menggunakan media sosial untuk mencari informasi (Marsono, 2014).

Kasus positif COVID-19 yang diumumkan pada awal Maret 2020, menjadi awal menyebarnya virus corona di Indonesia. Dengan adanya berita tentang kasus positif tersebut, masyarakat mulai mencari tahu tentang apapun yang berkenaan dengan virus corona, termasuk pencarian informasi melalui Twitter. Dengan tingginya keingintahuan masyarakat tentang virus corona tersebut, beberapa pihak mulai menyebarkan informasi dengan membuat akun khusus tentang COVID-19 di Indonesia. Selain akun – akun independen yang memberitakan tentang kasus COVID-19, akun akun twitter resmi milik pemerintah, baik pemerintah pusat maupun daerah juga terus menginformasikan mengenai kasus COVID-19 di Indonesia.

Meskipun tingkat keingintahuan masyarakat cukup tinggi pada awal masa pandemi, tetapi belum diketahui secara pasti tingkat partisipasi masyarakat terhadap twit atau cuitan yang terkait dengan informasi tentang COVID-19, baik pada awal terjadinya pandemi maupun ketika kondisi mulai memasuki Normal Baru.

Pengumpulan data dari sosial media atau social media mining menjadi salah satu cara untuk mengetahui tingkat partisipasi dari pengguna. Partisipasi dalam era teknologi dapat dianggap sebagai salah satu cara yang digunakan oleh masyarakat untuk mengutarakan opini atau sebagai masukan kepada pembuat kebijakan untuk pengambilan keputusan. Informasi tentang pandemi yang sedang terjadi perlu disebarkan seluas – luasnya sehingga masyarakat mendapatkan pengetahuan yang sebanyak – banyaknya dengan tujuan untuk bersinergi dengan pihak terkait dalam menurunkan tingkat penularan.

Dalam penelitian ini, dilakukan pengumpulan data dari seluruh twit yang dikirim oleh akun akun yang mengkhususkan diri untuk membahas COVID-19 di Indonesia serta twit dari akun resmi pemerintah. Tujuan dari penelitian ini adalah untuk mengetahui tingkat partisipasi masyarakat terhadap

informasi tentang COVID-19 di Indonesia yang disebarkan oleh akun independen maupun oleh akun resmi pemerintah, baik pemerintah pusat maupun daerah. Hipotesa awal yang menjadi dasar dari penelitian ini adalah adanya tingkat partisipasi yang cukup tinggi dari masyarakat tentang informasi terkait dengan COVID-19 di Indonesia. Kemudian, dari tingkat partisipasi yang cukup tinggi tersebut muncul pertanyaan mengenai besarnya tingkat partisipasi masyarakat terhadap informasi tentang COVID-19 pada saat sudah memasuki era kebiasaan baru yang notabene berlangsung setelah tiga bulan sejak kasus COVID-19 merebak di Indonesia. Dari pertanyaan – pertanyaan tersebut, penulis ingin mengetahui tingkat partisipasi masyarakat terhadap informasi yang disajikan terutama pada platform Twitter baik pada saat awal masa pandemi maupun pada saat memasuki era Kebiasaan Baru.

Penggunaan sosial media juga menjadi semakin penting dalam hal penyebaran berita. Meskipun televisi dan media cetak tetap menjadi sumber berita yang paling penting, tetapi penelitian yang melibatkan responden dari Denmark, Perancis, Jerman, Italia, Jepang, Spanyol dan Inggris menunjukkan bahwa peranan sosial media semakin penting dalam penyebaran berita.

Beberapa penelitian telah dilakukan dengan melibatkan data yang diambil dari Twitter. Di antaranya adalah penelitian tentang identifikasi dan prediksi *engagement* masyarakat dengan menggunakan data dari twit yang diambil dari akun twitter pemerintah. Penelitian ini dilakukan oleh Nur Siyam, dkk pada tahun 2019 dengan objek penelitian di jazirah Arab. Penelitian dilakukan dengan mengumpulkan 55,089 twet dalam periode satu tahun. Twit yang dikumpulkan memiliki karakteristik tipe, tanggal dan waktu. Metode yang digunakan untuk memprediksi *engagement* dari pengguna adalah dengan menggunakan metode machine learning yaitu Random Forest dan Adabost. Hasil yang didapatkan dari penelitian ini adalah type twit yang di-*post* mempengaruhi *engagement* dari pengguna. Pengaruh tertinggi terhadap tingkat *engagement* pengguna adalah pada twit dengan tipe video dan gambar. Hari pengiriman twit juga berpengaruh pada tingkat *engagement* pengguna. Dalam penelitian tersebut, didapatkan hasil bahwa jika twit dikirimkan pada hari kerja, maka pengaruh terhadap tingkat *engagement* pengguna meningkat. (Siyam, N, 2020).

Penelitian berikutnya meneliti tentang pengukuran tingkat *engagement* pengguna pada chat tentang kesehatan masyarakat. Penelitian dilakukan pada *chat* yang dilakukan oleh @TulanePRC dan @FitNOLA dengan menggunakan tagar #LiveFitNOLA. *Chat* yang dilakukan berfokus pada kesehatan di New Orleans. Chat tersebut digunakan untuk menyebarkan informasi dan mendapatkan partisipasi dari masyarakat dalam diskusi terbuka tentang budaya kesehatan di New Orleans.

Penelitian dilakukan dengan menggunakan metode Social Network Analysis dan Content Analysis. Penghitungan tingkat *engagement* dilakukan dengan menggunakan *Engagement Ratio* dan *Return On Engagement*.

Data yang digunakan diambil dari hasil chat yang dilakukan selama 75 menit pada tanggal 5 Maret 2015. Pada proses *chat* tersebut didapatkan 744 twit dari 66 peserta. Hasil yang didapatkan dari penelitian tersebut adalah rasio *engagement* dan ROE bergantung pada audiens, konteks pembicaraan, ruang lingkup, skala pembicaraan dan tujuan dari chat yang dilakukan di Twitter. (Rahman, Z, 2016)

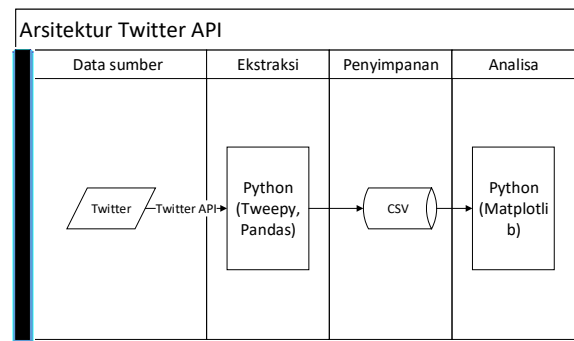
Pengukuran *engagement* pada pengguna sosial media banyak digunakan untuk melihat tingkat efektifitas sosial media. Tingkat efektifitas sosial media dapat dimanfaatkan oleh pemasar sehingga para pemasar dapat menentukan konten yang baik dan tepat untuk memasarkan produk mereka.

Penelitian dengan menggunakan data twitter tentang COVID-19 juga telah dilakukan oleh beberapa pihak. Di antaranya adalah penggunaan data twitter untuk analisa sentimen dan pola yang sering muncul. Data yang digunakan pada penelitian ini adalah 600.000 twit dengan tagar #COVID dan #coronavirus yang dikirimkan antara tanggal 27 Februari 2020 sampai dengan 25 Maret 2020. Analisa dilakukan pada twit berbahasa Inggris dengan menggunakan metode pendekatan lexicon-based. Kata – kata yang belum diolah kemudian dijadikan input beserta lexicon yang berisi daftar sentimen. Sentimen lexicon yang digunakan adalah AFINN, Bing dan NRC. Hasil dari penelitian tersebut menunjukkan bahwa sentimen negative terhadap COVID berada dalam level yang tinggi pada periode awal, kemudian naik pada minggu kedua dan menurun pada minggu – minggu berikut. Sebaliknya, sentimen positif yang berada pada level cukup rendah pada minggu pertama, menunjukkan kenaikan pada minggu – minggu berikutnya.

2. METODE PENELITIAN

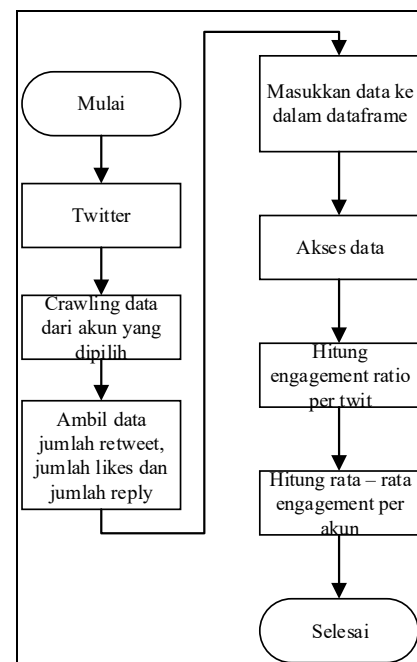
Pada penelitian ini, metode yang digunakan adalah dengan melakukan pengumpulan data dari twitter yang bersumber pada akun @KawalCOVID19, akun BNPB, akun Kementerian Kesehatan dan akun WHO Indonesia. Data twit dikumpulkan dengan menggunakan API Twitter yang dapat digunakan secara gratis. Proses *crawling* data menggunakan bahasa pemrograman Python dan Google Colaboratory sebagai editor. Banyak library berbasis bahasa Python yang dapat digunakan untuk mengakses API Twitter. Penulis menggunakan library Tweepy yang menyediakan fungsi – fungsi untuk mengakses API Twitter.

Pada gambar 1, ditunjukkan arsitektur dan alur proses pengumpulan data dengan menggunakan API dari Twitter.



Gambar 1. Arsitektur Twitter API

Penggunaan API Twitter dibatasi pada pengguna yang telah mendaftarkan sebagai developer pada twitter. Setiap pengguna yang telah mendaftar sebagai developer, akan mendapatkan kode akses tertentu setelah membuat aplikasi dan memberikan penjelasan tentang aplikasi yang akan dibuat kepada pihak twitter. Setiap akun developer tidak serta merta dapat menggunakan akunnya untuk mengakses API Twitter, karena Twitter juga membatasi hak akses pada setiap akun yang terdaftar. Misalnya, satu akun hanya bisa mengakses data twitter sampai dengan 7 hari ke belakang, atau bisa mengakses sampai dengan data 30 hari ke belakang.



Gambar 2. Alur penelitian

Pada gambar 2, ditunjukkan alur penelitian yang dilakukan. Penghitungan ratio engagement dilakukan dengan menggunakan rumus sebagai berikut:

$$\text{rasio engagement} = \frac{\left(\frac{\text{Interaksi}}{\text{jumlah twit}} \right) / \left(\frac{\text{Rata rata impression}}{\text{Rata – rata reach}} \right)}{\text{Rata – rata reach}} \times 100 \quad (1)$$

Dengan menggunakan rumus (1), maka langkah yang harus dilakukan setelah pengumpulan data tweet adalah menghitung nilai interaksi, rata – rata impression dan rata – rata reach untuk setiap tweet. Interaksi didapatkan dari jumlah retweet dan like untuk setiap tweet yang dikirimkan. Pada penelitian ini, difokuskan untuk menghitung rasio engagement pada setiap tweet. Berdasarkan riset yang dilakukan oleh Xu Han, penghitungan engagement pada tweet tunggal tetap menggunakan rumus (1), tetapi rata – rata impression/rata – rata reach dinilai 1.

Proses pengambilan data twitter menggunakan tools Tweepy yang menggunakan bahasa pemrograman python. Tweepy adalah library Python untuk mengakses Twitter API. Baris kode yang digunakan untuk mengakses Twitter API dengan menggunakan Tweepy adalah sebagai berikut:

```
auth = tweepy.OAuthHandler(consumer_key,
consumer_secret)
auth.set_access_token(access_token,
access_token_secret)
api = tweepy.API(auth)
```

Consumer Key, Consumer Secret, Access Token dan Access Token Secret diberikan oleh Twitter kepada pengguna yang mendaftarkan diri sebagai developer Twitter. Sedangkan kode yang digunakan untuk mengakses jumlah *retweet* dan jumlah *likes* adalah sebagai berikut:

```
tmpTweets = api.user_timeline(username)
tmpTweets.retweet_count
```

Dengan menggunakan kode `user_timeline` di atas, pengguna dapat memasukkan nama pengguna yang akan diambil datanya. Dalam penelitian ini, username yang dimasukkan adalah KawalCOVID19, BNPB dan username lain yang dipergunakan untuk penghitungan tingkat *engagement* pengguna.

API yang digunakan untuk mendapatkan jumlah retweet adalah `retweet_count`. Dengan menggunakan API tersebut, maka jumlah retweet dari sebuah tweet dapat dihitung. Penghitungan ini tidak termasuk pada jumlah quote tweet atau tweet dengan menggunakan komentar. Untuk menghitung jumlah likes pada sebuah tweet, dapat menggunakan kode `like_count`. Twitter menyediakan API tersebut untuk menghitung jumlah total like dari sebuah tweet, baik yang membayar maupun tidak.

Setelah mendapatkan keseluruhan tweet, jumlah likes dan jumlah retweet, maka hasil dari pengumpulan data tersebut disimpan dalam file csv. Setiap data yang didapatkan dari beberapa akun, disimpan dalam file dengan format csv tersendiri. Selanjutnya, file csv tersebut dibaca kembali dengan menggunakan kode python sehingga dapat dihitung rasio *engagement* dari keseluruhan tweet yang sudah didapatkan.

Beberapa tweet yang diambil tidak membahas tentang COVID, seperti tweet dari BNPB Indonesia, tidak semuanya berisi tentang COVID. Karena itu, setelah mendapatkan keseluruhan tweet pada rentang waktu yang ditentukan, sebelum dilakukan penghitungan rasio *engagement*, dilakukan filterisasi tweet. Filter yang dilakukan adalah dengan mengambil semua tweet yang mengandung kata COVID.

Proses filterisasi diawali dengan membaca file csv hasil dari pengumpulan data dengan menggunakan API twitter. Setiap hasil pengumpulan data yang sudah dimasukkan ke dalam file dengan format csv, dibaca dengan menggunakan library pandas yang dimiliki oleh bahasa pemrograman python. Hasil pembacaan kemudian divisualisasikan dengan memanfaatkan library matplotlib.

3. HASIL DAN PEMBAHASAN

Data yang didapatkan pada akun @KawalCOVID19 dari tanggal 20 April 2020 sampai dengan 20 Juni 2020 dituliskan pada tabel berikut:

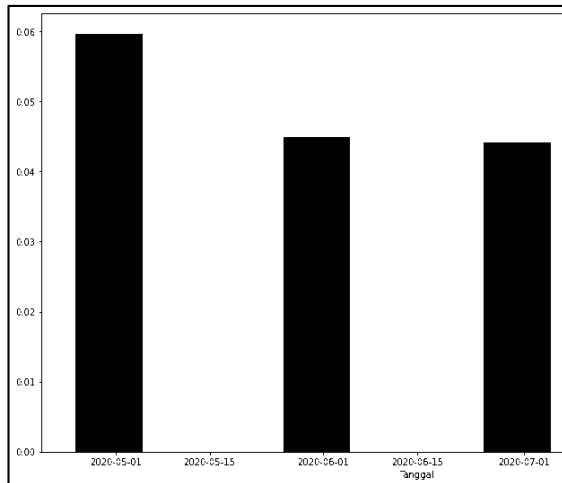
Tabel 1. Contoh hasil pada akun @KawalCOVID19

Tanggal	Teks	Jumlah Retweet	Jumlah Likes
2020-04-21 05:19:30	Hi Tweepers, \nhttps://t.co/cN9jguxwV 9 akan hadi...	14	11.0
2020-04-20 11:44:14	Mari saling bantu saat #COVID19indonesia https...	59	82.0
2020-04-20 09:02:13	Sebaran 185 kasus baru COVID-19 di Indonesia p...	132	145.0
2020-04-20 08:58:55	Indonesia mengumumkan 185 kasus baru #COVID19 ...	209	142.0
2020-04-20 00:41:19	♥□□□ https://t.co/08aSYNpfaN	22	53.0

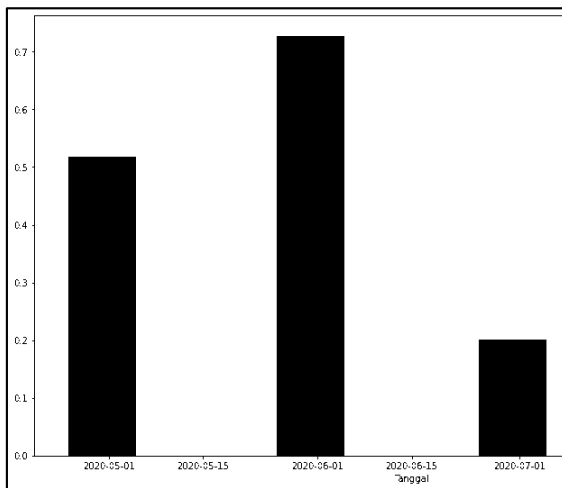
Pada tabel tersebut ditunjukkan jumlah data yang didapatkan, teks dari tweet yang dikirimkan, jumlah retweet dan jumlah likes untuk setiap tweet.

Untuk data yang didapatkan dari akun BNPB, harus difilter terlebih dahulu untuk mendapatkan teks tweet yang mengandung COVID. Jumlah data yang didapatkan pada tweet terkait COVID adalah 322 data. Filter terhadap teks tweet tersebut dilakukan pada semua akun yang tidak membahas COVID secara spesifik, seperti akun dari pemerintah provinsi Jawa Timur atau dari akun Kementerian Kesehatan.

Proses penghitungan dibagi menjadi 3, yaitu periode April, Mei dan Juni. Hasil penghitungan *engagement rate* pada setiap bulan ditunjukkan pada gambar 3 dan 4.



Gambar 3. Data rata – rata engagement rate BNPB



Gambar 4. Data rata – rata engagement rate akun KawalCOVID19

Dari gambar 3 dan 4 dapat dilihat bahwa terdapat penurunan tingkat *engagement* pada twit yang berhubungan dengan COVID. Pada data akun KawalCOVID19, tingkat *engagement* meningkat pada bulan kedua, yaitu periode Mei. Kemudian menurun drastis pada periode bulan ketiga, yaitu bulan Juni. Pada periode awal pandemi, yaitu bulan April dan Mei, masyarakat masih antusias untuk mencari informasi tentang virus corona yang sedang melanda Indonesia.

Dari gambar 3, ditunjukkan bahwa tingkat *engagement* pengguna pada bulan April sebesar 0,06 persen. Kemudian turun menjadi 0,044 persen pada bulan Mei dan turun kembali menjadi 0,043 persen pada bulan Juni. Sedangkan dari gambar 4, yaitu gambar yang menunjukkan tingkat rasio *engagement* dari KawalCOVID19, menunjukkan bahwa tingkat *engagement* pengguna adalah 0,5 persen pada bulan April, kemudian meningkat signifikan pada bulan Mei menjadi 0,7 persen dan menurun secara signifikan pada bulan Juni menjadi 0,2 persen.

Penurunan rasio *engagement* ditunjukkan oleh kedua gambar di atas pada bulan Juni. Penurunan rasio *engagement* memang tidak signifikan pada

akun BNPB, sedangkan pada akun KawalCOVID19, rasio *engagement* terlihat jelas tingkat penurunannya pada bulan Juni atau pada saat istilah New Normal diperkenalkan.

4. KESIMPULAN

Kesimpulan dari penelitian ini adalah adanya penurunan data tingkat *engagement* pada twit yang berhubungan dengan COVID. Hal ini tentu saja tidak dapat diartikan bahwa masyarakat mulai abai dengan berita tentang COVID, tetapi hal ini dapat dijadikan rujukan untuk meningkatkan upaya dalam penyebaran informasi tentang COVID di Indonesia. Dengan hasil yang didapatkan, maka dapat dilakukan langkah – langkah lain untuk meningkatkan ketertarikan masyarakat akan berita tentang COVID. Untuk pengembangan dari penelitian ini, dapat dilakukan penelitian tentang hubungan antara konten atau isi dari berita yang disebarkan dengan tingkat *engagement* pengguna.

DAFTAR PUSTAKA

- ABD-ALRAZAQ, A., ALHUWAIL, D., HOUSEH, M., HAMDI, M., & SHAH, Z. (2020). Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. *J Med Internet Res*, 22.
- BARBOUR, K. (2020). Using Twitter to study audience engagement with characters and actors. In M. Griffiths, & K. Barbour, *Making Publics, Making Places* (pp. 180 - 192). Adelaide: University of Adelaide Press.
- BOUDJELIDA, A., MELLOULI, S., & LEE, J. (2016). Electronic Citizens Participation: Systematic Review. *ICEGOV '15-16: Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance* (pp. 31 - 39). New York: ACM.
- DATABOKS KATADATA. (2019). Retrieved 2020, from Datastories: <https://databoks.katadata.co.id/datapublish/2019/02/08/berapa-pengguna-media-sosial-indonesia>
- JAYANI, D. H. (2020). *Databoks Katadata*. Retrieved July 2020, from Data Stories: <https://databoks.katadata.co.id/datapublish/2020/02/26/10-media-sosial-yang-paling-sering-digunakan-di-indonesia>
- MARSONO, C. (2014). *What Social Medias Are Indonesians Using? (2014 Data)*. Retrieved 2020, from linkedin: <https://www.linkedin.com/pulse/what-social-medias-indonesians-using-2014-data-chandra-marsono?redirectFromSplash=true>

- NIELSEN, R. K., & SCHRØDER, K. C. (2014). The Relative Importance of Social Media for Accessing, Finding, and Engaging with News. *Digital Journalism*, 472-489.
- RAHMAN, Z., SUBERAMANIAN, K., ZANUDDIN, H., MOGHAVVEMI, S., & NAZIR, M. H. (2016). Social Media Engagement Metrics Analysis - "Study on Fan Page Content". *researchagate.net*, 71-76.
- ROSENBERG, H., SYED, S., & REZAIE, S. (2020). The Twitter pandemic: The critical role of Twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. *CJEM*, 1-4.
- SIYAM, N., ALQARYUTI, O., & ABDALLAH, S. (2020). Mining government tweets to identify and predict citizens engagement. *Technology in Society*, 1-10.
- SLIPP, N. (2019). Using Tweets as a Measure of Sponsor Engagement: Halifax Pride 2018. *DJIM*.
- VOORVELD, H. A., NOORT, G., MUNTINGA, D. G., & BRONNER, F. (2018). Engagement with Social Media and Social Media. *Journal of Advertising*, 38-54.