



Analisis Penggunaan Data Sosial Media Untuk Penelitian: Studi Etika Dan Regulasi Informatika

Sani Shalihamidiq¹, Nana Mulyana Maghfur², Rahmat Alfanza³, Didi Juardi⁴

^{1,2,3,4}Universitas Singaperbangsa Karawang

Abstract

Received: 13 Agustus 2022

Revised: 19 Agustus 2022

Accepted: 25 Agustus 2022

Public data contained on social media is widely used by researchers, the data can be in the form of text, photos, videos to locations and hashtags created by social media users themselves. Many researchers use social media as a place to find data for free because the data is already public. However, ethically the use of data violates existing rules, because many social media users do not know that their data is used as research material. In this article, the author aims to analyze the ethics of researchers in using data sourced from public social media, as well as regulatory provisions regulated by the government and service providers on the use of data.

Keywords: Ethic, Internet, Research, Social Media

(*) Corresponding Author: sani.shalihamidiq18059@student.unsika.ac.id

How to Cite: Shalihamidiq, S., Maghfur, N., Alfanza, R., & Juardi, D. (2022). Analysis of Social Media Data Usage for Research: Ethical Studies and Informatics Regulation. *Jurnal Ilmiah Wahana Pendidikan*, 8(17), 610-616. <https://doi.org/10.5281/zenodo.7081580>

INTRODUCTION

Media sosial adalah platform daring yang seringkali bersifat mobile dimana fungsi utamanya yaitu untuk mendukung pembuatan dan pertukaran konten dari satu pengguna kepada pengguna lain (Kaplan & Haenlein, 2010). Dalam dua dekade terakhir kemajuan teknologi telah mendorong cepatnya perluasan akses internet di banyak negara dimana jumlah pengguna media sosial juga telah mengalami perkembangan pesat menjadi sekitar 4,2 miliar orang pengguna aktif di seluruh dunia (Kemp, 2021). Hal ini membuat media sosial menjadi tempat dimana banyak macam data terkumpul dan dapat diakses siapa saja selama konten yang diunggah pengguna bersifat publik.

Karena ketersediaan data yang banyak, beragam, serta kemudahan akses inilah banyak penelitian saat ini dilakukan dengan mengambil data pada media sosial. Sebagai contoh, penelitian belandaskan pada data hasil web scraping dari media sosial twitter guna melakukan perangkuman suatu topik diskusi telah beberapa kali dilakukan pada Bahasa Inggris maupun beberapa bahasa lainnya (Kemaloglu & Küçükşille, 2018; Li & Zhang, 2020). Salah satu tren yang signifikan dilakukan yaitu penggunaan alat otomatis berupa program untuk melakukan proses mining pada data media sosial untuk memudahkan pengambilan data dalam jumlah banyak sekaligus dimana data digunakan untuk memprediksi perilaku pengguna media sosial dimasa depan.

Selain penjabaran diatas, sejumlah manfaat potensial penggunaan data media sosial dalam penelitian telah dijelaskan dalam literatur, termasuk kemampuan untuk menjangkau lebih banyak peserta daripada yang mungkin dilakukan (Moorhead et al., 2013), mampu menganalisis tren dan asosiasi dalam



kumpulan besar data akses terbuka (Paul & Dredze, 2011), mengurangi biaya melakukan penelitian dalam populasi besar (Munson et al., 2013), dan banyak potensi lainnya yang tidak bisa disebutkan satu persatu.

Terlepas dari kelebihan-kelebihan yang sudah disebutkan, oleh karena kompleksitas interaksi antara individu, kelompok, dan sistem teknis di ruang daring ini menghadirkan sejumlah tantangan bagi akademisi yang ingin menggunakan data media sosial dalam penelitian dimana tantangan terbesar yaitu dari segi etika penggunaan data (Berry, 2004; Eysenbach & Till, 2001). Seperti persepsi yang bervariasi dan batasan yang tidak jelas antara ruang publik dan pribadi, serta kesulitan memastikan anonimitas dan menjaga privasi subjek data yang identitasnya tidak boleh disamarkan atau dapat dengan mudah disimpulkan dari postingan dan afiliasi mereka (Taylor & Pagliari, 2018).

Selain karena pengguna media sosial yang masih kurang memahami soal jaminan privasi dan keamanan datanya di media sosial, belum adanya peraturan dan regulasi yang jelas dari pemerintah, khususnya pemerintah Indonesia jika dibandingkan dengan negara lain, menyebabkan kondisi dimana banyak peneliti memiliki dilemma etis atas data yang digunakannya dalam penelitian. Oleh karena itu, melalui tulisan ini, penulis bertujuan untuk menganalisa apakah penelitian – penelitian yang telah dilakukan dengan menggunakan data yang bersumber dari media social sudah sesuai dengan etika penggunaan data yang ada dengan tidak melanggar privasi dari perspektif pengguna juga tidak melanggar hak kekayaan intelektual.

METHODS

Dalam melakukan penelitian ini kami mengambil beberapa artikel penelitian yang dalam prosesnya menggunakan data publik bersumber dari media sosial. Selanjutnya artikel-artikel tersebut kami analisis berdasarkan pandangan pengguna media sosial yang datanya dipakai serta kami juga menganalisis apakah penggunaan data dalam penelitian sudah sesuai dengan regulasi pemerintah dan juga regulasi pada penyedia layanan.

RESULTS & DISCUSSION

Result

1. Pandangan pengguna atas penggunaan data media sosial mereka

Berdasarkan beberapa hasil studi diketahui bahwa sebagian besar pengguna merasa tidak aman jika data yang mereka hasilkan dari penggunaan media sosial secara berkala, digunakan untuk tujuan lain seperti rekomendasi iklan (contohnya pada kasus *Facebook* dimana dari sampel pengguna 60 % menyatakan keberatan) hingga analisis perilaku atau sentiment terhadap suatu topik meski dalam beberapa studi lain pengguna lebih banyak menunjukkan sikap acuh terhadap keterbukaan data mereka dibandingkan merasa terancam (Kennedy et al., 2017).

Meski masih amat sulit untuk menentukan batasan antara yang termasuk data privasi atau bukan, namun secara definisi dimana data privasi merupakan data yang dapat menunjukkan identitas atau karakteristik khas yang hanya dimiliki oleh seseorang diketahui bahwa yang bisa dikategorikan sebagai privasi yaitu data seperti biodata, lokasi terkini, kontak, serta nomor – nomor kredensial akses (contoh: pin atm, password akun, dan lain - lain) (Iyengar, 2002).

2. Peraturan pemerintah yang mengatur perihal penggunaan data media sosial Pemerintah Indonesia belum secara spesifik memiliki kebijakan atau regulasi mengenai perlindungan data pribadi dalam satu undang-undang khusus, dimana peraturan yang ada saat ini masih terpisah dan tersebar di beberapa undang-undang dan hanya mencerminkan aspek perlindungan data pribadi secara umum (Oktavira, 2020). Salah satu yang paling mendekati yaitu terdapat pada UU ITE **Pasal 26 ayat (1) UU 19/2016** yang berbunyi :

*“Kecuali ditentukan lain oleh peraturan perundang-undangan, penggunaan setiap informasi melalui media elektronik yang menyangkut **data pribadi seseorang** harus dilakukan atas persetujuan orang yang bersangkutan.”*

UU ITE ini sendiri masih sangat kontroversial karena banyak ditemukan pasal multitafsir dan banyak pihak meminta agar UU ini dikaji ulang.

DISCUSSION

Tabel 1 menampilkan pembahasan mengenai keetisan penggunaan data dari beberapa penelitian yang menggunakan data bersumber dari media sosial. Nilai keetisan diukur berdasarkan seberapa besar kemungkinan privasi pengguna terekspos dan apakah penggunaan data sudah sesuai peraturan yang berlaku.

Judul	Penulis	Sum ber Data	Penggun aan Data	Keetis an
<i>Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis</i>	(Budiharto & Meiliana, 2018)	<i>Twitter</i> pengguna media sosial	Penulis menggunakan data dari banyak user twitter karena menggunakan tagar untuk mencari data yang sesuai dan data lalu dianalisis sentimennya berdasarkan kata tidaknya kata berkonotasi negatif atau positif.	Penelitian ini cukup etis dimana meski data banyak digunakan, namun adanya proses penghilangan kata yang merujuk ke pengguna (@user) membuat penelitian ini menjaga data yang dianalisis tetap anonim.
<i>Twitter Scrapping for</i>	(Herlati et al., 2020)	<i>Twitter</i> pengguna tertentu	Penulis hanya menggunakan	Penelitian ini sudah sangat etis

Profiling Education Staff	pada media sosial <i>Twitter</i>	data yang berasal tweet akun pribadinya sendiri saja dan studi kasus menyesuaikan tempat dimana penulis bekerja.	dikarenakan penulis peduli pada persoalan privasi dari proses penelitian ini, sehingga yang digunakan hanya data pribadinya saja, bukan data orang lain.	
Penerapan Data Mining Untuk Klasifikasi Perilaku Pengguna Twitter Terhadap 5 Gempa Bumi Terbesar di Indonesia Tahun 2019 Ditinjau Dari Model Psikologi Dynamic Disaster Model Dan Rumor Theory	(Bendesa, 2021)	<i>Tweet</i> pengguna media sosial <i>Twitter</i>	Data <i>tweet</i> pengguna lebih dahulu dianalisa secara manual oleh peneliti dan psikolog sebelum akhirnya dianalisa menggunakan komputer melalui metode <i>machine learning</i> .	Penelitian ini dirasa kurang etis dimana berbeda dengan penelitian – penelitaian diatas yang langsung melakukan analisa menggunakan komputer sehingga perspektif peneliti tidak ikut terbawa, penelitian ini melakukan analisa manual terlebih yang membuat peneliti dapat mengetahui kondisi psikologis pengguna dimana pengguna

				tersebut belum tentu setuju untuk hasil ketikannya ditafsirkan oleh orang lain.
<i>Hatenography On Twitter During the Covid-19 Pandemic in Indonesia: Hate Speech Case Against Anies Baswedan</i>	(Bajari et al., 2021)	<i>Tweet dan reply</i> pada akun pengguna tertentu media sosial <i>Twitter</i>	Data dari <i>tweet</i> dan <i>reply</i> pada 6 akun yang dipilih untuk merepresentasikan golongan pro dan kontra pemerintah dianalisis sentimennya menggunakan metode <i>machine learning</i> .	Penelitian ini sulit untuk ditentukan keetisannya karena 6 akun yang digunakan belum diketahui apakah akun pribadi peneliti atau bukan, jika iya maka penelitian ini sudah lumayan etis dan jika tidak maka penelitian ini kurang etis.
<i>Analisis Sentimen dan Klasifikasi Respon Pengguna Media Sosial Facebook Terhadap Pendidikan Daring Pada Masa Pandemi Covid-19 di Indonesia</i>	(Rizal, 2021)	<i>Post</i> pengguna media sosial <i>Facebook</i>	Data <i>post</i> <i>Facebook</i> yang merupakan hasil <i>web crawling</i> digunakan untuk pembuatan model menggunakan algoritma <i>Naïve Bayes</i> dimana telah dilakukan <i>preprocessing</i> standar pengolahan	Penelitian ini meski masih kekurangan detail tentang keetisan penggunaan datanya, masih tergolong cukup etis karena data yang didapat langsung diolah menggunakan

teks sebelumnya	n komputer, sehingga mengurangi bias perspektif peneliti.
--------------------	--

CONCLUSION

Dari hasil pembahasan penelitian – penelitian yang dikaji keetisannya, masih ada beberapa masalah kurang etisnya penelitian yang muncul baik dari workflow penelitian sampai pemilihan data yang kurang diperhatikan keetisannya, meski beberapa peneliti sudah sadar akan pentingnya privasi dari data yang digunakan. Maka dari itu perlu dilakukan sosialisasi dan regulasi ulang baik dari sisi pemerintah maupun dari sisi penyedia platform sehingga peneliti – peneliti yang belum menyadari adanya masalah etika pada penelitiannya dapat dengan sadar memperbaiki diri agar kedepannya peneliti bisa lebih berhati – hati dan tidak menyepelekan perihal etika penggunaan data terutama pada data yang didapat dari media sosial.

REFERENCES

- Bajari, A., Istiqomah, R. N., Koswara, I., & Erlandia, D. R. (2021). Hatnography On Twitter During the Covid-19 Pandemic in Indonesia: Hate Speech Case Against Anies Baswedan. *Review of International Geographical Education Online*, 11(5), 68–78.
- Bendesa, I. W. P. (2021). *Penerapan Data Mining Untuk Klasifikasi Perilaku Pengguna Twitter Terhadap 5 Gempa Bumi Terbesar di Indonesia Tahun 2019 Ditinjau Dari Model Psikologi Dynamic Disaster Model Dan Rumor Theory* [Skripsi, UNIVERSITAS AIRLANGGA]. <https://doi.org/10.10.10.%20LAMPIRAN%20.pdf>
- Berry, D. M. (2004). Internet research: Privacy, ethics and alienation: an open source approach. *Internet Res.* <https://doi.org/10.1108/10662240410555333>
- Budiharto, W., & Meiliana, M. (2018). Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis. *Journal of Big Data*, 5(1), 51. <https://doi.org/10.1186/s40537-018-0164-1>
- Eysenbach, G., & Till, J. E. (2001). Ethical issues in qualitative research on internet communities. *BMJ*, 323(7321), 1103–1105. <https://doi.org/10.1136/bmj.323.7321.1103>
- Herlawati, H., Trias Handayanto, R., Ekawati, I., Meutia, K. I., Asian, J., & Aditiawarman, U. (2020). Twitter Scrapping for Profiling Education Staff. *2020 Fifth International Conference on Informatics and Computing (ICIC)*, 1–6. <https://doi.org/10.1109/ICIC50835.2020.9288607>
- Iyengar, V. S. (2002). Transforming data to satisfy privacy constraints. *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 279–288. <https://doi.org/10.1145/775047.775089>

- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59–68. <https://doi.org/10.1016/j.bushor.2009.09.003>
- Kemaloglu, N., & Küçüksille, E. U. (2018). Automatic Text Summarization Methods Used on Twitter. *Data Science and Applications*, 1(1), 9–15.
- Kemp, S. (2021, January 27). *Digital 2021. We Are Social UK*. <https://wearesocial.com/uk/blog/2021/01/digital-2021-uk/>
- Kennedy, H., Elgesem, D., & Miguel, C. (2017). On fairness: User perspectives on social media data mining. *Convergence*, 23(3), 270–288. <https://doi.org/10.1177/1354856515592507>
- Li, Q., & Zhang, Q. (2020). Abstractive Event Summarization on Twitter. *Companion Proceedings of the Web Conference 2020*, 22–23. <https://doi.org/10.1145/3366424.3382678>
- Moorhead, S. A., Hazlett, D. E., Harrison, L., Carroll, J. K., Irwin, A., & Hoving, C. (2013). A New Dimension of Health Care: Systematic Review of the Uses, Benefits, and Limitations of Social Media for Health Communication. *Journal of Medical Internet Research*, 15(4), e1933. <https://doi.org/10.2196/jmir.1933>
- Munson, S. A., Cavusoglu, H., Frisch, L., & Fels, S. (2013). Sociotechnical Challenges and Progress in Using Social Media for Health. *Journal of Medical Internet Research*, 15(10), e226. <https://doi.org/10.2196/jmir.2792>
- Oktavira, B. A. (2020, August 4). *Dasar Hukum Perlindungan Data Pribadi Pengguna Internet—Klinik Hukumonline*. [hukumonline.com. https://hukumonline.com/klinik/detail/ulasan/lt4f235fec78736/dasar-hukum-perlindungan-data-pribadi-pengguna-internet/](https://hukumonline.com/klinik/detail/ulasan/lt4f235fec78736/dasar-hukum-perlindungan-data-pribadi-pengguna-internet/)
- Paul, M., & Dredze, M. (2011). You Are What You Tweet: Analyzing Twitter for Public Health. *Proceedings of the International AAAI Conference on Web and Social Media*, 5(1), 265–272.
- Rizal, A. (2021). *Analisis Sentimen dan Klasifikasi Respon Pengguna Media Sosial Facebook Terhadap Pendidikan Daring Pada Masa Pandemi Covid-19 di Indonesia* [S1_i, STMIK Widya Cipta Dharma]. <https://repository.wicida.ac.id/3728/>
- Taylor, J., & Pagliari, C. (2018). Mining social media data: How are research sponsors and researchers addressing the ethical challenges? *Research Ethics*, 14(2), 1–39. <https://doi.org/10.1177/1747016117738559>