User acquisition and profile of COVID-19's health education website: A descriptive study

Avinindita Nura Lestari,[1*] Tiara Putri Leksono,[2] Reyfal Khaidar,[3] Ekky Novriza Alam,[4] and Lutfan Lazuardi[5]

[1]Faculty of Medicine, Universitas Islam Bandung, Bandung, Indonesia
[2]Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia
[3]Faculty of Medicine, Universitas Islam Negeri Syarif Hidayatullah Jakarta, Jakarta, Indonesia
[4]Information System Department School of Industrial Engineering, Telkom University Bandung, Indonesia
[5]Department of Health Policy and Management, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Indonesia

ABSTRACT A health start-up in Bandung, Indonesia made initiatives to educate people about COVID-19 prevention through downloadable scripts and audio in the form of Public Service Announcements provided in 19 local languages through the website called HEUP or Health Empowerment and Educational Project. This study aimed to know the characteristics of profile users accessing the HEUP website containing health promotion material through a descriptive observational approach. Data came from Google Analytics which collects traffic from the main website. We examined the audience data, consisting of demographics and geographical distribution. Additionally, we observed the acquisition data which helped us see the website traffic. A significant difference was found in this study in the age group, while the gender group did not have any substantial difference, with only 8% disparity. By geographical distribution, 60% of top users were located in cities, especially in Java Island. Direct traffic, interestingly, made up almost 86% of all traffic. Twitter ranked at the top for the social media traffic in our case. In conclusion, it is necessary to promote credible information in COVID-19 preventive measures and help maintain the accessibility of information.

1. Introduction

On December 31st, 2020, the World Health Organization (WHO) China Country Office reported a pneumonia of unknown cause from Wuhan city, China.[1] A week later, the etiology was identified as a new type of Coronavirus, named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses (ICTV). Later, it was called Coronavirus Disease 2019 (COVID-19) by the WHO on February 11, 2020.[2]

People are exploring health information by seeking the information via the Internet. They seek information from various websites disseminating health information online, with likely consequences for the health care system.[3] The Internet is a large source of health information and has the capacity to influence its users. Another phenomenon that exists is the digital divide, which has also been studied in Indonesia. From a survey by Lim in 2011, and Wattegama and Soehardjo in previous years, 85% of the respondents owned a mobile phone and only 27% of the mobile phone owners accessed the Internet.[4] It showed that not all of Indonesia's population used the gadget effectively. Surprisingly, less than half of our population used the Internet. Based on Global Connectivity Index, Indonesia had a population of 260 million according to the United Nations, and 93.4 million of whom were Internet users.[5] Meanwhile,

*Correspondence: avininditanl@gmail.com
Faculty of Medicine, Universitas Islam Bandung, Jalan Tamansari No.20, Bandung, Jawa Barat, Indonesia, 40116

over time, the Association of Internet Service Providers in Indonesia (APJII) 2018 survey showed that the 2018 national Internet penetration rate was 64.8%, showing growth from the previous year of 54.68%.[6]

Digital divide is not only defined by the matter of Internet accessibility, or we could say, a single dimension. There are more variables to it. Some dimensions that we should look for include: 1) the subjects (e.g. divide between individuals, countries, etc.); 2) different characteristics among the subjects (e.g. income, geography, age, etc.); 3) the ways of connecting (mere access or effective adoption), and 4) tools (e.g. phones, Internet, digital-TV, etc.).[7] To know how important information was received in a community, we considered collecting demographic data and data traffic were of primary importance. By collecting these data, we could see if individual digital divide occurs, such as gender, age, and geography. Thus, our target population were people who were willing and had access to visit the HEUP or Health Empowerment and Educational Project website for improving their knowledge on COVID-19 health promotion.

Reflecting on the content of the health messages, it needs to convey accurate information in a way that is understood by the public, particularly the COVID-19 preventive measures, since there is currently no vaccine or specific antiviral treatment. Additionally, the prevention behaviors noted in this study could be applicable worldwide.[8]

While there has been a lot of information spreading in social media, television, radio, and we were fortunate enough to have the access through our fingertips. This is not usually the case with underprivileged people with little access to technology. In Latin America, the current tariff structure has an inhibiting effect on service consumption by the poor[8], whereas they also have the same risk from being infected or of infecting people with COVID-19 just as we do. Uniquely in Indonesia, in every street corner there are many mosques/small churches/local gathering places with their loudspeaker on. HEUP, a health startup based in Bandung, saw this as an opportunity and made initiatives to educate people about COVID-19

prevention through downloadable scripts and audio in the form of Public Service Announcements (PSA). As per June 2020, HEUProject had released the downloadable scripts and PSA in Bahasa and 19 different local languages.[9] We saw the PSA as a good instrument for health promotion because the WHO has also pointed out ways of bridging the language divide in health, highlighting that language could be a barrier to accessing relevant and high-quality information, and for delivering appropriate health care.[10]

Accordingly, reflecting on the multiple dimensions of the digital divide, this study aimed to know the characteristics of profile users accessing the website for HEUP's downloadable educational script regarding COVID-19 via demographic data and website traffic.

2. Method

This study used a descriptive observational design conducted during the Sounds of Nusantara Project by the HEUProject. The study was conducted using secondary data. The data were taken from Google Analytics as a third party application that helped record traffic of HEUProject's website from 23rd March-7th April 2020 to get descriptive characteristics of users assessing the website at covid19.heuproject.com. The data used consisted of demographics (age and gender) and geographical distribution, particularly cities. Not only examining the audience data, we also observed the acquisition data that helped us in seeing website traffic. There were no ethical concerns since the data inputs were all anonymous.

The data were divided into several categories according to the purpose of the analysis, including user data, location, reference sources, and social media reference sources. The purpose of the analysis stage was to find a correlation between user users, location and reference sources that lead users to access the website. Google Analytics can help filter data as needed, such as filtering website visitor data from social media (for example Twitter, Instagram, and Facebook). The results of the screening process are shown in Table 3. The process and analysis results are described in the next section.

Table 1. The top 10 list of Users based on Cities

Group	User detail		
	Users	New users	Session(s)
Total	9,788	9,504	12,444
Jakarta	2705	2559	3375
Bandung	1861	1761	2377
Depok	832	790	1010
Surabaya	630	605	747
Makassar	420	409	521
Batam	247	221	292
Semarang	225	212	292
Medan	197	184	234
Surakarta	193	181	225
Brighton	190	190	190

3. Result

By April 7th, 2020, 2 weeks after the launching of the downloadable script and PSA, according to Table 1, there were 9,788 users who assessed the website distributed over 231 cities worldwide. The top cities were mainly located in Java Island. Surprisingly, the traffic also reached out to another city outside of Indonesia, which was Brighton.

Regarding Figure 1, based on the data we acquired, the percentage of men (54%) that accessed the web of COVID-19 education was more than women (46%). Moreover, the data showed that the primary age group that accessed this site was in the range of 25-35 years old (33.5%), followed by groups of range 18-24 years old (27,5%), and then by 35-44 years old (15.5%). A small percentage under 10% was identified in both groups of ages 55-65 and >65 years old.

As depicted in Figure 2, the age group distribution of users was calculated by 12,444 sessions. Table 2 displays information about the ways viewers reached the website, or in this case, the website traffic. Among 9,788 users, as many as 8,425 users directly accessed the web, while a significant amount of nearly a thousand users accessed through social media platforms. Website referral access was done by 351 users and the rest were done by organic search. Among 977 users, as presented in Table 3, Twitter traffic reached the top rank by 569 users and 689 sessions.

4. Discussion

4.1 Geographical distribution by cities

In terms of users, based on data from the Association of Internet Service Providers in Indonesia (APJII), Indonesia's Internet penetration in urban areas has a figure of 74.1% for Internet users and 25.9% for Internet non-users. From the same survey, in all regions of Indonesia, the highest contribution of Internet use was from Java as much as 55.7%, followed by Sumatra as much as 21.6%. Nearly 11% of users are spread across Sulawesi-Maluku-Papua and no more than 7% for each in Kalimantan, Bali and Nusa Tenggara. There was a significant difference when compared to rural areas, where Internet users are around 61.6% and not internet users are 38.4%.[6] This was reflected in the results of the study shown in Table 1, where 6 out of top 10 cities were from Java Island in Indonesia. Surprisingly, Makassar was in the 5th rank of top cities by the number of users.

The setting and population in Indonesia, which is an archipelago country with numerous ethnic and linguistic groups make the health promotion system become more complex and trickier in terms of health promotion messages. Additionally, the political decentralization structure in the health system can affect the method of health promotion.[11] We found that our study also illustrates these differences and thus it serves as a lesson learnt for such a case of digital divide.

4.2 Age and gender

Our study results show that the most frequent age group that accessed the HEUP site was the range of 25-35 years old, which was stated previously in a research by Hargittai showing significant differences in online skills, particularly by age, and also described that people in their teens and people in their 20s are quicker than people in their 30s and 40s.[12] Moreover, a study of 605 respondents by Puspitasari and Ishii in Indonesia in 2015 found that more educated and younger people accessed the Internet more often although the study was centered in big cities and did not include geographical (rural-urban) variables in the survey distribution.[4]

Gender
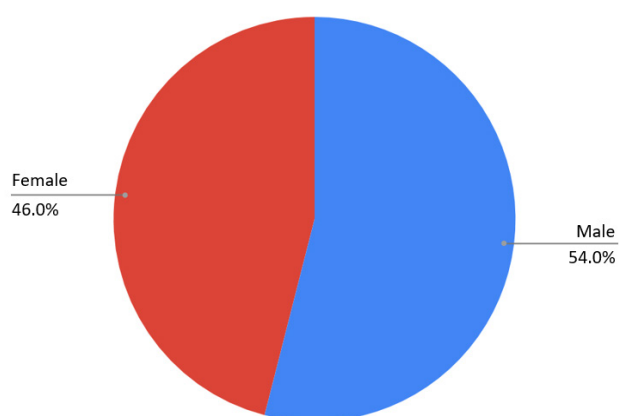
Figure 1. Gender distribution of 12,444 user sessions

Age group
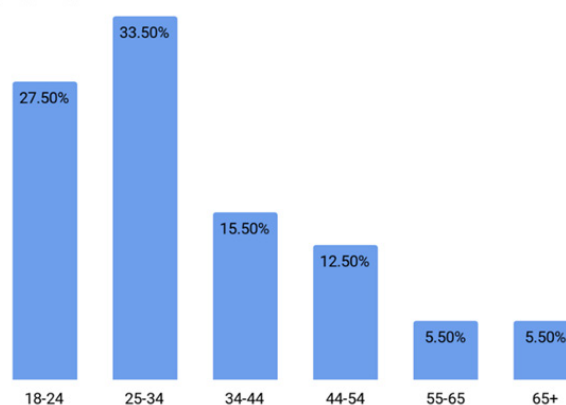
Figure 2. Age group distribution of 12,444 user sessions

Table 2. Website traffic generated by user

	Number of users
Direct	8,425
Social	977
Referral	351
Organic search	66
Total	9,788

Table 3. Social media traffic of the website

Social Network	**Users**	**New Users**	**Sessions**
Twitter	569	555	689
Instagram Stories	200	195	207
Instagram	138	133	146
Facebook	72	70	80
YouTube	1	0	2
Total	977	953	1,124

The study findings of the gender group did not show any significant difference, with only 8% disparity. This is consistent with the research that showed there is no influence of gender on whether people are able to efficiently navigate the content of the Web and how long they take to do so. On average in comparison, women completed 4.19 tasks compared to men's 4.26 average success rate. The average total time spent on the five tasks for women was 14.6 minutes whereas for men it was 12.9 minutes. Neither of these differences were statistically significant, suggesting that there is no influence of gender on whether people are able to efficiently navigate the content of the Web and how long they take to do so.[13]

Reflecting on the importance of young people as Internet users, it is also stated that Indonesia Broadband Plan includes a national digital literacy program to improve Internet and communication technology (ICT) literacy on the national level to promote adoption and meaningful use of broadband.[13] Recently, however, the Indonesian

government removed the school subject of information technology from the new curriculum in 2013.[14] Promoting ICT skills may be the key to narrowing the digital divides that exist in Indonesia. It is recommended that future policy should consider the advantages of improving ICT literacy.

4.3 Website traffic

There are several methods of website acquisition conducted by visitors that created traffic. Direct traffic comes when visitors manually enter the website URL or have the page bookmarked and this traffic made up most of the users' traffic. In our case, it made up 86% of all traffic. Another different type is referral traffic. The difference lies in that these come from recommendations from a different site other than search engines for a specific website. As a user clicks on the hyperlink in a website that further takes them to a different page of the website, it is considered as the referral visit.[15] Another result is generated from organic search. The term 'organic search' encompasses a search generating results without

paid advertisements. The results of this search are listed on the search engine result page and appear when they are relevant to the keywords since they are generated by popularity and common usage.[16]

The utilization of referral sources can also be based on the category of the website. Social media websites can generate traffic, called the social traffic. These kinds of traffic, interestingly, made up almost 26 percent of all traffic. The utilization of referral sources can also be based on the category of the website. When websites use social media such as Facebook, Twitter, YouTube, and other similar sources to pull in traffic, the traffic generated is regarded as social traffic.[16] In this study, we can see that Twitter ranked at the top for social media traffic.

The online conversation about health is made possible by two factors. The first one is the availability of social tools, and the second one is the motivation among people to connect.[17] Turning to Twitter and COVID-19, there is a relevant illustration from a prior infodemiological study in South Korea in February 2020 that collected Coronavirus Twitter data from 43,382 users and 78,233 conversations. The study found that the spread of information among people who used the word 'Coronavirus' was faster. This highlighted the positive role of individuals and groups that further directed public attention to the pandemic crisis. Tweets containing medically framed news articles were found to be more popular than tweets that included news articles adopting non-medical frames.[18]

Our study reflects that there are still disparities in accessing information between cities located in the island of Java and the rest of the country. To engage a wider audience, such important information, especially in the terms of public health promotion, should be collectively broadcasted. We found data that showed the age group of 25-35 years old has the potential to reach out to more people about the content of information.

5. Conclusions

Most users accessing the website were 25-35 years old and located in Java Island. This implies that individual digital divide exists among the users. The way which the users assessed the information

was via direct traffic or entering the URL manually and Twitter ranked the highest for social media traffic. Dissemination of information was lacking in organic search aspects. This means that it should be necessary for any health information provider to ensure good promotion and user engagement as well as the content to ensure the information can be disseminated widely. The website creator could take measures on how to increase search engine optimization reflecting on the low level of organic search traffic. We also can see the potential of the 25-35 years old age group to promote the information on the Internet. By promoting these messages, this can help maintain and increase the accessibility of information.

Acknowledgement

Conflict of interests

The author(s) declared no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

References

1. Kementrian Kesehatan. Pedoman Pencegahan dan Pengendalian Coronavirus Disease (COVID-19). Pedoman Pencegah dan Pengendali Coronavirus Dis. 2020;4(Revisi ke-4):1–125

2. Pan L, Mu M, Yang P, Sun Y, Wang R, Yan J, et al. Clinical characteristics of COVID-19 patients with digestive symptoms in Hubei, China. *Am J Gastroenterol*. 2020;115(5):766–73.

3. Siswanta. Informasi Kesehatan di Media Online. *Jurnal Ilmu Komunikasi UNISRI Surakarta*. 2015;13(3):210-223.

4. Puspitasari L, Ishii K. Digital divides and mobile Internet in Indonesia: impact of smartphones. *Telemat Informatics*. 2016;

5. Global Connectivity Index: Country profile Indonesia [Internet]. www.huawei.com. 2019 [cited 18 June 2020]. Available from: https://www.huawei.com/minisite/gci/en/country-profile-id.htm

6. Asosiasi Penyelenggara Jasa Internet Indonesia. Laporan survei penetrasi & profil perilaku pengguna internet di Indonesia 2018. Asosiasi Penyelenggara Jasa Internet Indonesia. 2018.

7. Hilbert M. The end justifies the definition: the manifold outlooks on the digital divide and their practical usefulness for policy-making. *Telecomm Policy*. 2011;35(8):715

8. Basch CH, Hillyer GC, Meleo-Erwin ZC, Jaime C, Mohlman J, Basch CE. Preventive behaviors conveyed on YouTube to mitigate transmission of COVID-19: cross-sectional study. *JMIR Public Heal Surveill*. 2020;

9. Project HE and E. No Title [Internet]. 2020 [cited 2020 Jun 11]. p. 1. Available from: https://www.talenthouse.com/i/1839/submission/424125/38c866e1

10. Bridging the language divide in health [Internet]. World Health Organization. 2015 [cited 18 June 2020]. Available from: http://dx.doi.org/10.2471/BLT.15.020615

11. Mahendradhata Y, Trisnantoro L, Listyadewi S, Soewondo P, Marthias T, Harimurti P et al. The Republic of Indonesia health system review. *Health Systems in Transition*. 2017;7(1):1-8.

12. Hargittai E. Second-level digital divide: differences in people's online skills. FM [Internet]. 2002Apr.1 [cited 2020Jun.18];7(4). Available from: https://journals.uic.edu/ojs/index.php/fm/article/view/942

13. Lukito D Tuwo. Rencana pembangunan broadband nasional. 2013;(September).

14. Vota W. Should Indonesia be teaching ICT in schools? - ICTworks [Internet]. ICTworks. 2014 [cited 18 June 2020]. Available from: https://www.ictworks.org/should-indonesia-be-teaching-ict-in-schools/

15. Chakrabortty K, Jose E. Relationship analysis between website traffic, domain age and Google indexed pages of e-commerce websites. *IIM Kozhikode Soc Manag Rev*. 2018;7(2):171–7.

16. Duklan M, Bahuguna H. Search engine optimization techniques for organic search : a theoretical approach. *Uttaranchal Bus Rev*. 2016;(April):45–55.

17. Fox S. Health Topics: 80% of Internet users look for health information online. Pew Internet & American Life Project. 2011.

18. Park HW, Park S, Chong M. Conversations and medical news frames on Twitter: infodemiological study on COVID-19 in South Korea. *J Med Internet Res*. 2020.