

A temporal analysis of depression related tweets - a case study in Finland

Jonas Tana¹, Bao Cai² and Leonardo Espinosa-Leal²

¹Arcada University of Applied Sciences, Graduate School and Research, Helsinki, Finland jonas.tana@arcada.fi

² Arcada University of Applied Sciences, Graduate School and Research, Helsinki, Finland

Abstract

Depression is one of the most burdensome diseases in the world. A problem that depression presents, is the fact that it is connected with a high rate of unwillingness to seek professional help, and therefore many aspects of depression go unreported, affecting our understanding of it. Nowadays, individuals turn to online platforms for help and support, which creates vast amounts of data. This infodemiological study utilised data from Twitter to identify temporal patterns of behaviours related to depression in Finland. The findings of this study can be used to improve the impact of public health measures in relation to depression.

Keywords

depression, mental health, social media, twitter, infodemiology

1 INTRODUCTION

Mental health related health challenges are on the rise in today's society. One of these challenges, depression, which is one of the most common mental health diseases, has been projected to become the leading cause of disability in the next 20 years. In fact, it has already been ranked as one of the most burdensome diseases in the world by the World Health Organization (WHO) [1]. According to the WHO, depression is both a symptom and an illness, and characterized by sadness, loss of interest or pleasure, feelings of guilt or low self-worth, disturbed sleep or appetite, feelings of tiredness and a poor ability to concentrate [2]. Not only can depression be debilitating and cause impairments in an individual's everyday life, depression can also affect chronic health conditions, such as cancer, diabetes and obesity, as well as cardiovascular diseases [1]. Depression typically has a broad spectrum and can range from anxiety and milder forms to episodic or chronic severe depression. Depression can also be cyclic in its occurrence, and factors like temporal aspects and natural rhythms, such as the shift between night and day or the seasonal changes, can influence an individual's tendency to suffer from depression and depression like symptoms [3, 4 pp. 325-326, 5, 6 pp. 1-3, 7]. These kind of periodic variations in depression, and more generally in mood and mental health are familiar for most individuals, both at a circadian level (24-hour cycles), circaseptan (weekly cycles) and on a seasonal level (yearly cycles), especially in the southern and northern hemispheres, where changes in external conditions are evident. Reasons for these variations in mood are, apart from the environmental changes, likely endocrinological as well as social [8, 4 pp. 325-326]. However well known, from a public health perspective, there is a challenge in monitoring these variations. A reason for this is that more traditional methods, such as surveys, are gathered at a specific time

once a year or even less often, and thus lack the ability to identify trends, changes, and variations in health status within smaller time periods and intervals [1]. Depression incidence is also usually studied based on data from registers in healthcare institutions and organisations [9]. A challenge in relation to this, is that depression is connected with a high rate of unwillingness to seek professional help and can therefore go unreported. Whereas the first step toward appropriate treatment would be to seek professional help, individuals have been shown more likely to turn to online platforms and resources for help because of the stigma and barriers to care associated with depression [10, 11, 12, 13]. This presents a large gap in our knowledge in relation to depression and its variations. These large temporal gaps can hinder the development of effective and timely intervention programs and delay public health officials from acting on changes in health status [1, 9]. In the case of depression this is highly relevant, as a delay in initial treatment contact is associated with worse outcomes, despite effective interventions [12, 14].

The use of online generated user data may provide a method to overcome some of these gaps. As already mentioned, people suffering from depression are likely to turn to online platforms, and today, this online health information behaviour is extensive [11]. The vast amounts of data this generates, in search engines, on websites and in social media, not only allows us to study behaviours in relation to depression that have previously gone unreported, but also makes it possible to capture them in real-time and at a fine-grained temporal scale [1]. The use of data from online communication to study health related behaviours has been on the rise in the past decades, and has been called infodemiology, which can be defined as *the science of distribution of and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health*

The 18th Scandinavian Conference on Health informatics, Tromsø, Norway, August 22-24, 2022. Organized by UiT The Arctic University of Norway. Conference Proceedings published by Linköping University Electronic Press at <https://doi.org/10.3384/ecp187>. © The Author(s). This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>

and public policy [15, p. 155]. The underlying assumption in infodemiology is, that in the age of the Internet, there is a relationship between population health on one hand, and information and communication patterns on the internet, on the other [16]. Therefore, changes in peoples' health status or behaviour, as well as public attention and attitudes are echoed in changes in online information and communication patterns [15]. The infodemiology approach has been effectively utilised to study temporal variation of online health behaviour in relation to depression, both in search engines and on discussion forums [11, 17]. Especially social media, where individuals are increasingly sharing their thoughts, emotions, and health concerns, can serve as a resourceful medium for mining information about the public's mental health and behaviours in a non-invasive manner on a large scale over time and space [13, 18].

Of all social media platforms, Twitter, the microblogging service, has been the most utilised platform for conducting research on expressed emotions [18]. Tweets have in previous research been analysed in relation to a myriad of health-related topics, ranging from influenza outbreaks and vaccinations to obesity, physical activity and drinking problems [19]. Mental health related tweets have also been analysed in various studies, and previous studies have shown, that people post about their depression and their treatment for depression on Twitter [1]. In relation to temporal variations, the expression of emotions on Twitter have also been studied, with different aims, ranging from the detection of emotional contagion and change in public opinions to identifying mental disorders and measuring population mood before, during and after natural disasters [18]. A few previous studies have found circadian and seasonal patterns in the content of emotionally loaded wordings in Twitter messages. However, the majority of studies have focused on simple frequency analysis, or content analysis as well as classification by machine learning approaches, and disregarded the temporal variations and patterns [19, 20]. Thus, these previous investigations do not examine the full potential of the platform. Moreover, to the authors' best knowledge, no studies to date have analysed depression related tweets limited to Finland. According to Official Statistics of Finland [21], in 2017, 11 percent of the population aged 15 or over used Twitter. In early 2022, Twitter was reported to have 759.5 thousand users in Finland, which would indicate that 15.8 percent of the population aged 15 or above in Finland used Twitter in early 2022 [22]. Even if the amount of Twitter users in Finland is limited, and only comprises of a small segment of the adult population, it can provide useful information, and complement previous infodemiological studies, on depression related online health information behaviour in Finland. Moreover, Finland, with its northern location and extreme variations in external conditions such as seasonal temperature and daylight, is an interesting subject to study health related phenomena from a temporal aspect [23, 24].

Therefore, the aim of this preliminary study is to identify temporal patterns and variation, as well as periods of heightened interest, in mental health topics, in this case depression, on Twitter in Finland.

2 METHOD

For this study, we collected and analysed timestamped and geo-located tweets identified with hashtags and terms relating to depression in Finnish (a list of included words can be found in Appendix 1). The terms were identified via the Google Trends top related query terms function, which resulted in 35 terms. It is worth noting that, even though the language and geographic location were set to Finnish and Finland, respectively, some of the keywords used in the search were English, simply because it is not uncommon to mix languages, especially with scientific vocabulary involved. We limited the tweets to include only tweets that originated in Finland geographically. UTC time was converted to the local time of the country in which tweets originated (Finland), after which we computed Hour and Day as numeric variables. We used the European convention to order the days of the week (leading to values of 0 for Sunday and 6 for Saturday).

Tweets were collected via the Twitter API version 2, specifically the `/2/tweet/search/all` endpoint of the Twitter API. The study utilised a subset of features the endpoint provided. A query is the main search feature containing the keywords, geographic location, and the language setting for the tweets. In addition, the date range was used to specify tweets' creation time, tweet, and users' fields to control the amount of data received and expansions to bypass the tweet's shortened form. A script was made to programmatically collect the tweets which contained any word in a set of keywords and were posted during a certain timeframe. The script accepts, at minimum, plain text keywords or a keyword text file, start and end time, and the location to output the file containing the tweets. In addition, geographic location, language preference, and data fields can also be passed to the script to control the receiving contents.

Aside from the text and the ids associated with each data point, the timestamp when the data point was created, the type of the data point, and the author id were also recorded. No personally identifiable information (PII) was collected in this study. The time frame of the data collection was from 2017 to 2021, including three categories: *quoted*, *replied_to*, and *tweet*. The complete set included 5114 data points, but we only studied direct tweets; therefore, we are limited to 2751 data points in this study. In the plots, the data is normalised yearly, and the average is calculated over the total number of years, five in this paper.

Statistical calculations and data analysis was performed using python programming language version 3.8.13, including the libraries Pandas (version 1.2.3) and Scipy (version 1.5.2). Plots were done using Matplotlib (version 3.5.1).

3 RESULTS

The A total of 2751 ($n=2751$) tweets were identified containing one or more of the chosen keywords in relation to depression during the chose five-year period, between 2017 and 2021. Out of sample of 2751 tweets, we found 1030 unique users. Of these unique users, for instance, 690 users posted one tweet, 172 users posted two tweets, and 49 users posted three tweets. Therefore, most of the unique users (999) posted less than 10 tweets in the dataset. We found one user that posted 249 tweets. Statistical

significance of the data distributions depicted in Figures 2-5 were obtained using the standard Kolmogorov–Smirnov test available in the Scipy python package. The obtained p-values are consigned in Table 1. All results reject the null hypothesis ($p < 0.05$).

Year	Monthly (Fig. 2)	Yearly (Fig. 3)	Weekly (Fig. 4)	Hourly (Fig. 5)
2017	0.0015	1.26e-12	0.022	7.36e-06
2018	0.0017	1.40e-12	0.023	4.34e-06
2019	0.0016	1.41e-12	0.024	7.44e-06
2020	0.0018	7.49e-13	0.018	4.33e-06
2021	0.0016	7.85e-13	0.018	4.26e-06

Table 1. Kolmogorov-Smirnov p-values for each series of data per year and distribution as presented in Figures 2-5.

As can be seen in Figure 1, the yearly distribution of depression related tweets shows a peak in 2019, with 678 tweets. In 2021 again, the number of tweets mentioning depression is at its lowest, with 388 tweets.

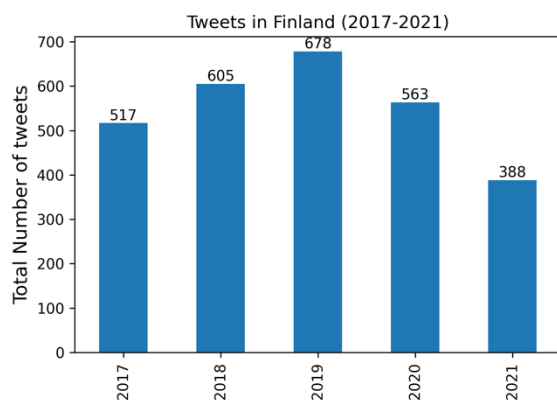


Figure 1. Yearly distribution of depression related tweets in Finland between 2017-2021.

On a seasonal scale, our analysis shows that tweets that mention depression follow a bimodal curve, with higher activity during spring and autumn. On an average level, the autumn peak is more significant compared to the spring peak. However, as can be seen in Figure 2, there is an exception for the year 2020, where a clear spring peak can be identified, and where activity is significantly higher during February. This February peak in 2020 is followed by a high activity in the end of the year 2019, specifically in November 2019.

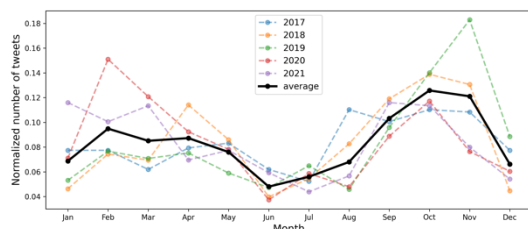


Figure 2. Monthly distribution of tweets that mention depression in Finland during 2017-2021.

A more detailed analysis of the seasonal variations on a weekly level, as illustrated in Figure 3, shows that the November peak in 2019 is located more towards the end of November, and the February 2020 peak again, is located in mid-February, shortly after the WHO declared a public

health emergency of international concern in relation to the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that was rapidly spreading worldwide.

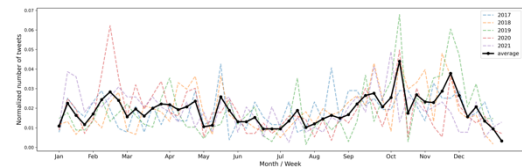


Figure 3. Yearly distribution of tweets mentioning depression on a weekly scale in Finland for the years 2017-2021.

On a circaseptan, or weekly, scale, tweets containing references to depression related terms, show a higher distribution towards the beginning of the week, reaching a small peak on Wednesday and Thursday, from where onwards activity rapidly starts to decline. The lowest number of tweets that mention depression related terms can be displayed on Sunday. The weekly distribution of tweets is presented in Figure 4.

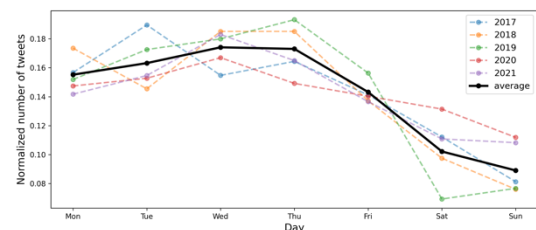


Figure 4. Weekly distribution of tweets containing depression related terms for the years 2017-2021.

The diurnal variation of tweets that contain depression related content is presented in Figure 5. Our analysis reveals a clear unimodal curve, where activity starts to rise in early morning from 03:00 onwards and reaching a peak at between 07:00 and 08:00 in the morning. Activity then starts to decline in the evening, from 19:00 onwards. Activity then is at its lowest between 22:00 and 03:00.

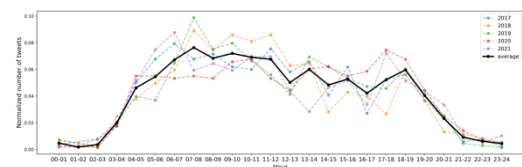


Figure 5. Distribution of tweets that contain depression related content on a diurnal (24-hour) scale for the years 2017-2021.

4 DISCUSSION

Our results indicate that tweets in Finnish containing depression related content follow temporal patterns, on a seasonal (yearly), circaseptan (weekly), as well as circadian (24-hour) level. The bimodal curve identified for tweets that contain depression related content in Finnish is similar

to previous infodemiological studies in relation to depression in Finland. Peaks in spring and autumn have previously been identified for search engine [24] and discussion forum activity [11] in relation to depression. Moreover, these seasonal findings, with peaks during spring and autumn, also correspond with previous findings of seasonality in diagnosed depression, hospital admissions and suicide rates in the Scandinavian countries and more broadly in the Northern Hemisphere [11]. The late November peak in activity is also similar to Dzogang et al. [8], who found that negative affect in messages on Twitter is significantly periodic and over-expressed in late November. The significantly higher peak in activity in depression related tweets during February 2020 identified in this study is most likely related to the Covid-19 pandemic and does validate previous findings that have indicated a significant increase in mental health symptoms early in the pandemic [25]. A more detailed content analysis of the tweets for these peak time-periods is already planned, with the aim to identify reasons for heightened activity.

On a circaseptan level, the findings in this study are somewhat contradictory to previous findings of online health behaviour in relation to depression in Finland, where peaks have been identified during Sundays and Mondays [11, 24]. Sundays and Mondays have also been identified as days when lower mood or depression symptoms are more prevalent [26]. From a broader perspective, health contemplations in general have been shown to peak in the beginning of the week and decline towards the end of the week [14]. Moreover, Twitter usage has in some studies been shown to decline towards weekends, which could explain the weekend trough in activity [27, 28].

As for the diurnal variations, activity on Twitter reveals a different temporal pattern compared to search engine queries in relation to depression in Finland, where significant peaks are visible during night-time, and troughs in activity during daytime [17]. The same night-time peaks have also been found in discussion forums for discussions related to depression [11]. However, a similar early morning peak to the one identified in this study, in tweets containing words representing negative affect has been identified by Golder & Macy [29] globally, in four English speaking regions. They identified a peak between 03:00 and 04:00 in the morning. Moreover, Ten Thij et al. [30] found, that people suffering from depression showed higher levels of “rumination” and “self-reflection” between 03:00 and 06:00 in the morning. The rising activity during early-morning, between 03:00 and 06:00, could be an indicator for disturbed sleep quality and sleep patterns, something that is a common symptom for depression [2].

An explanation for the differing results on a circaseptan as well as diurnal level, compared to the temporal patterns exhibited in search engines and discussion forums in Finland, could be the somewhat different use of Twitter compared to search engines and discussion forums. On Twitter, the largest user group consists of professional or business users, often more active during office hours, while the smallest user group is personal users. On Twitter, the personal, or casual users, are also the users that have a low or mild social engagement [31].

As can be seen, our findings indicate that collecting tweets offers an empirically based, objective and valid way of identifying and revealing temporal patterns and periods of heightened activity on Twitter in relation to depression. Data from Twitter can therefore help create a broader picture of online behaviours in relation to depression. The findings in this study also shows that comprehensive trend analysis of data from social media can reveal important insights that can be useful for timing more effective interventions and disseminate credible information related to depression [19]. For instance, detecting a rise in levels of depression related content could trigger social media platforms, in this case Twitter, to recommend internet-based intervention services, such as online cognitive behavioural therapy, to the users [13]. This would allow for early or timely optimal interventions, which have been shown to be of relevance in treatment response [14]. Moreover, this kind of infodemiological research can be extended to almost any health condition, and by collecting and analysing publicly available data, in near-real time, public health officials can be provided with early warning signals. This again can help provide important information to intervene by designing and implementing timely public health campaigns [15, 18]. It however needs to be noted, that this method of data collection and analysis are not meant to be a replacement for traditional public health surveillance systems or clinical diagnoses of depression, but as a compliment to them. As a substantial part of the population act on health-related issues without the involvement of health professionals, utilizing novel ways to detect these behaviours can complement more traditional data gathering methods, and aid in taking appropriate and needful public health measures [1, 15, 32].

5 LIMITATIONS

Our study does present some limitations. Firstly, there is no way to ensure that all tweets mentioning depression or related terms are strictly related to depression. A solution to this would be better filtering of the data and techniques, such as sentiment analysis to further refine the sample, could improve precision [19]. However, as the aim of this preliminary study is merely to identify temporal patterns in depression related online behaviour, future studies will utilize more precise tools to conduct further analysis on content of the tweets. Another limitation is the amount of Twitter users in Finland. With a relatively small percentage of users, the results cannot be generalised to a wider population, and the population observed on Twitter may display activity at times that are not corresponding with a broader population. Nevertheless, as more and more people are starting to engage in health-related behaviour on different online platforms, there is a need to examine these to be able to gain a holistic perspective on these behaviours. This again, is necessary in order to provide effective interventions, in this case for people that might need help with mental health related issues.

6 CONCLUSIONS

The aim of this study was to identify temporal variations and patterns of tweets containing depression related content in Finland. This preliminary study is the first of its kind in Finland to analyse the temporal variations and patterns of tweets that relate to depression. The results are somewhat

contradictory to previous findings from similar infodemiology studies, which advocates more research within this field, to gain a broad perspective of online behaviours in relation to depression. Generally, research on harnessing online data for understanding online health behaviours is still in its infancy, and as so often, more research is needed to harness the full potential that the analysis of this relatively new data generated on online media platforms can offer to tackle public health concerns [19]. The amount of online engagement in matters that relate to health keep rising on existing, as well as new platforms, which results in more and more user generated data. Therefore, it is necessary to pursue research on online platforms in order to improve representativeness of health behaviours, and as a result of that, the impact of public health measures.

7 REFERENCES

- [1] De Choudhury, M., Counts, S., Horvitz, E. "Social media as a measurement tool of depression in populations." In *Proceedings of the 5th annual ACM web science conference*, pp. 47-56. 2013.
- [2] Marcus, M., Yasamy, M. T., van Ommeren, M., Chisholm, D., Saxena, S. "Depression: A global public health concern." WHO Department of Mental Health and Substance Abuse, Vol. 1, pp. 6-8. 2012.
- [3] Antypa, N., Vogelzangs, N., Meesters, Y., Schoevers, R., Penninx, B. W. "Chronotype associations with depression and anxiety disorders in a large cohort study." *Depression and Anxiety*, Vol. 33, Issue 1, pp. 75-83. 2016.
- [4] Dunlap J.C., Loros J.J., DeCoursey P.J. (eds.). "Chronobiology: biological timekeeping." Sunderland MA: Sinauer Associates, 2004.
- [5] Maes, M., Meltzer, H. Y., Suy, E., De Meyer, F. "Seasonality in severity of depression: relationships to suicide and homicide occurrence." *Acta Psychiatrica Scandinavica*, Vol. 88, Issue 3, pp. 156-161. 1993.
- [6] Roenneberg, T. "Internal time: Chronotypes, social jet lag, and why you're so tired." Harvard University Press. 2012.
- [7] WHO. "Depression fact sheet." 2018. Available at: <https://www.who.int/news-room/fact-sheets/detail/depression> > [Accessed 02 May 2022].
- [8] Dzogang, F., Lansdall-Welfare, T., Cristianini, N. "Seasonal fluctuations in collective mood revealed by Wikipedia searches and Twitter posts." In *2016 IEEE 16th International Conference on Data Mining Workshops*, pp. 931-937. 2016.
- [9] Luo, J., Du, J., Tao, C., Xu, H., Zhang, Y. "Exploring temporal patterns of suicidal behavior on Twitter." In *2018 IEEE International Conference on Healthcare Informatics Workshop*, pp. 55-56. 2018.
- [10] Oliphant, T. "The information practices of people living with depression: constructing credibility and authority." Doctoral dissertation, The School of Graduate and Postdoctoral Studies, University of Western Ontario, London, Ontario, Canada, 2010.
- [11] Tana, J., Eirola, E., Eriksson-Backa, K. "Exploring temporal variations of depression-related health information behaviour in a discussion forum: the case of Suomi24." *Information Research*, Vol 25, Issue 2, paper 854. 2020.
- [12] Yoshikawa, E., Taniguchi, T., Nakamura-Taira, N., Ishiguro, S., Matsumura, H. "Factors associated with unwillingness to seek professional help for depression: a web-based survey." *BMC Research Notes*, Vol. 10, Issue 1, pp.1-6. 2017
- [13] Zhang, Y., Lyu, H., Liu, Y., Zhang, X., Wang, Y., Luo, J. "Monitoring depression trends on Twitter during the COVID-19 pandemic: Observational study." *JMIR infodemiology*, Vol 1, Issue 1, e26769. 2021.
- [14] Ayers, J. W., Althouse, B. M., Johnson, M., Dredze, M., Cohen, J. E. "What's the healthiest day?: Circaseptan (weekly) rhythms in healthy considerations." *American Journal of Preventive Medicine*, Vol 47, Issue 1, pp. 73-76. 2014a.
- [15] Eysenbach, G. "Infodemiology and infoveillance: tracking online health information and cyberbehavior for public health." *American Journal of Preventive Medicine*, Vol 40, Issue 5, pp. 154-S158. 2011.
- [16] Eysenbach, G. "Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet." *Journal of Medical Internet Research*, Vol. 11, Issue 1, e11. 2009.
- [17] Tana, J., Kettunen, J., Eirola, E., Paakkonen, H. "Diurnal variations of depression-related health information seeking: case study in Finland using Google Trends data." *JMIR Mental Health*, Vol. 5, Issue 2, e43. 2018.
- [18] Mayor, E., Bietti, L. M. "Twitter, time and emotions." *Royal Society open science*, Vol. 8, 201900. 2021.
- [19] McClellan, C., Ali, M. M., Mutter, R., Kroutil, L., Landwehr, J. "Using social media to monitor mental health discussions— evidence from Twitter." *Journal of the American Medical Informatics Association*, Vol 24, Issue 3, pp. 496-502. 2017.
- [20] Lampos, V., Lansdall-Welfare, T., Araya, R., Cristianini, N. "Analysing mood patterns in the United Kingdom through Twitter content." *arXiv preprint*, arXiv:1304.5507. 2013.
- [21] Official Statistics of Finland (OSF): "Using social media service, those aged 10 or over (2017) %, in Participation in leisure activities" [online]. Helsinki: Statistics Finland. Available at: http://www.stat.fi/til/vpa/meta_en.html > [Accessed 02 May 2022].
- [22] Kemp, S. "Digital 2022: Finland." [online]. 2022. Available at: <https://datareportal.com/reports/digital-2022-finland> > [Accessed 03 May 2022].
- [23] Basnet, S., Merikanto, I., Lahti, T., Männistö, S., Laatikainen, T., Vartiainen, E. and Partonen, T. "Seasonal variations in mood and behavior associate with common chronic diseases and symptoms in a population-based study." *Psychiatry Research*, Vol. 238, pp. 181-188. 2016.

- [24] Tana, J. "An infodemiological study using search engine query data to explore the temporal variations of depression in Finland." *Finnish Journal of eHealth and eWelfare*, Vol. 10, Issue 1, pp. 133-142. 2018.
- [25] Robinson, E., Sutin, A. R., Daly, M., Jones, A. "A systematic review and meta-analysis of longitudinal cohort studies comparing mental health before versus during the COVID-19 pandemic in 2020." *Journal of affective disorders*, Vol. 296, pp. 567-576. 2022.
- [26] Tumen, S., Zeydanli, T. "Day-of-the-week effects in subjective wellbeing: does selectivity matter?." *Social Indicators Research*, Vol. 119, Issue 1, pp. 139-162. 2014.
- [27] García-Palomares, J. C., Salas-Olmedo, M. H., Moya-Gomez, B., Condeco-Melhorado, A., Gutierrez, J. "City dynamics through Twitter: Relationships between land use and spatiotemporal demographics." *Cities*, Vol. 72, pp. 310-319. 2018.
- [28] Longley, P. A., Adnan, M., Lansley, G. "The geotemporal demographics of Twitter usage." *Environment and Planning*, Vol. 47, Issue 2, pp. 465-484. 2015.
- [29] Golder, S. A., Macy, M. W. "Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures." *Science*, Vol. 333, Issue 6051, pp. 1878-1881. 2011.
- [30] Ten Thij, M., Bathina, K., Rutter, L. A., Lorenzo-Luaces, L., van de Leemput, I. A., Scheffer, M. Bollen, J. "Depression alters the circadian pattern of online activity." *Scientific reports*, Vol. 10, Issue 1, pp. 1-10. 2020.
- [31] Uddin MM, Imran M, Sajjad H. "Understanding types of users on Twitter." arXiv preprint arXiv:1406.1335. 2014.
- [32] Ayers, J. W., Althouse, B. M., & Dredze, M. "Could behavioral medicine lead the web data revolution?" *Jama*, Vol. 311, Issue 14, pp. 1399-1400. 2014b.

Appendix 1 – terms related to depression used to identify tweets.

Masennus (eng. depression)

masennus oireet (eng. depression symptoms)

masennustesti (eng. depression test)

masennus testi (eng. depression test)

synnytyksen jälkeinen masennus (eng. postpartum depression)

masennus hoito (eng. depression treatment)

ahdistus (eng. anxiety)

keskivaikea masennus (eng. moderate depression)

psykoottinen masennus (eng. psychotic depression)

lapsen masennus (eng. childrens depression)

vakava masennus (eng. severe depression)

depression (eng. depression)

vaikea masennus (eng. serious depression)

nuorten masennus (eng. youth depression)

raskaus masennus (eng. pregnancy depression)

väsymys (eng. fatigue)

lievä masennus (eng. mild depression)

masennuslääkkeet (eng. depression medication)

mielenterveys (eng. mental health)

nuoren masennus (eng. adolescent depression)

masennus blogi (eng. depression blog)

itsemurha (eng. suicide)

masennus keskustelu (eng. depression discussion)

psykoosi (eng. psychosis)

masennuksen hoito (eng. depression treatment)

masennus itsehoito (eng. depression self

krooninen masennus (eng. chronic depression)

kaksisuuntainen mielialahäiriö (eng. bipolar disorder)

depression test

depression symptoms

manic depression

postpartum depression

crippling depression

clinical depression

high functioning depression