An analysis of Sentiments of Tweets about Patents

Alptekin Durmuşoğlu
Associate Professor
Department of Industrial Engineering
Gaziantep University, Gaziantep, Turkey
durmusoglu@gantep.edu.tr

Zeynep Didem Unutmaz Durmuşoğlu
Associate Professor
Department of Industrial Engineering
Gaziantep University, Gaziantep, Turkey
unutmaz@gantep.edu.tr

Abstract

Patents are documents about inventions that provide legal protection from which we can learn about technological development. Firms hold the monopoly of producing the relevant technology for a certain period of time, and in order to protect their technological superiority, they prepare the invention notification forms and complete their official applications to the patent office. The number of patents is a strong indicator of the R&D activities of a firm, country or region. For this reason, it is possible to write a large number of scientific articles that specifically analyze a technology or a region. Thus, what does society think of these official monopoly documents? One of the ways to find out what the society thinks about a subject is to follow the flow on the social media channel twitter. This study presents a partial analysis of what the public is talking about on twitter about patents. The tweets written in English, collected for a total of 2 months (106415 tweets), are subjected to a content analysis. When the hashtags of the tweets were examined, it was determined that the patents of some technologies (VR, crypto, blockchain, vaccine) were much trendy. In addition, sentiment analysis of tweet contents was also performed. It has been determined that while the emotional state is neutral in tweets about crypto patents, the average emotional intensity of tweets about vaccines is negative. In addition, it has been revealed through analyzes that these results differ according to countries and gender.

Keywords:
Sentiment analysis, patents, tweets, technologies, vaccines

1. Introduction

Today, due to the widespread use of the internet, many people express their opinions through social media services. Social media platforms such as Twitter, which are also available to ordinary members of the public, can provide unique information about public reactions to certain events and changes in their minds over time. It has been widely recognized that Twitter is a strong indicator of public awareness on a range of current social and medical issues (Crannell et al. 2016). Increases or decreases in tweet volume can also indicate which event is getting the most attention.

On the other hand, patents are documents that provide legal protection for the protection of the production right of technological innovations. It is common for patents to be analyzed to gain insight into the direction of technological development (Durmuşoğlu 2017; Durmusoglu and Unutmaz Durmusoglu 2021). However, it is noteworthy that no study has been conducted to understand what patents mean as technical documents for a non-technical audience. Technology can have positive and negative connotations for societies, as well as the perception of patents, which are official registrations of technologies, can be classified as positive or negative. How much has patents/technologies been a matter in Twitter, a channel of daily communication, and what emotional load do these tweets have? This study collects tweets about the patents in English within a certain date range and aims to analyze the key topics and the emotional load in the content of the tweets. Twitter's retrospective listing of tweets containing search terms for no more than one week required the query to be repeated daily. As part of the data collection process, all tweets that

might be relevant were programmatically searched using Webometric_Analyst_4.1 and removed from Twitter. This extensive search for tweets was conducted between 09/08/2021 and 23/11/2021. As a result, a total of 106415 English tweets containing the word "patent" were collected and the data was saved as a csv file. Remaining data and sentiment analysis were conducted using the free software Mozdeh.

2. Literature Review

The importance of user-generated social media content is increasing day by day. Automatically determining the emotional load of the content produced can play an important role in keeping the public's pulse on a subject. The term "emotion" represents a positive/negative opinion, emotion, feeling or thought expressed by an emotion holder (user). In general, sentiment analysis aims to automatically extract these sentiments from text. Sentiment analysis aims to examine textual features to automatically search for a sentiment at the word, sentence or document level (Bibi et al. 2022). Sentiment analysis in Twitter tackles the problem of analyzing the tweets in terms of the opinion they Express(Giachanou and Crestani 2016). Sentiment analysis is popular today in various fields such as public health monitoring (Paul et al. 2015), signal in venture capital financing (Tumasjan et al. 2021), election trends (Jungherr 2016), exploring public opinions on climate change (Wei et al. 2021), measuring the effect of public health campaigns (Ahmed et al. 2018).

3. Results and Discussion

The graphical representation of the time series is mainly used to create a graph of the volume of tweets on the topic over time. As demonstrated in Figure 1, although the number of tweets about patents has fluctuated, apart from the significant drop in October 2021, patents have always been a hot topic in social media. A total of 106415 tweets were created by 62383 different users. Most tweets were sent from bot accounts. Chief among these is the "patesalo" account, which shared the news of Japanese intellectual property with 1819 tweets. The second most tweeted account belongs to UK-based intellectual property consultancy OPUS IP. The third ranked account is a Microsoft boot account with 977 tweets. Since the purpose of this study is to analyze tweets of individuals, we excluded bot and corporate user accounts that posted more than three patents.

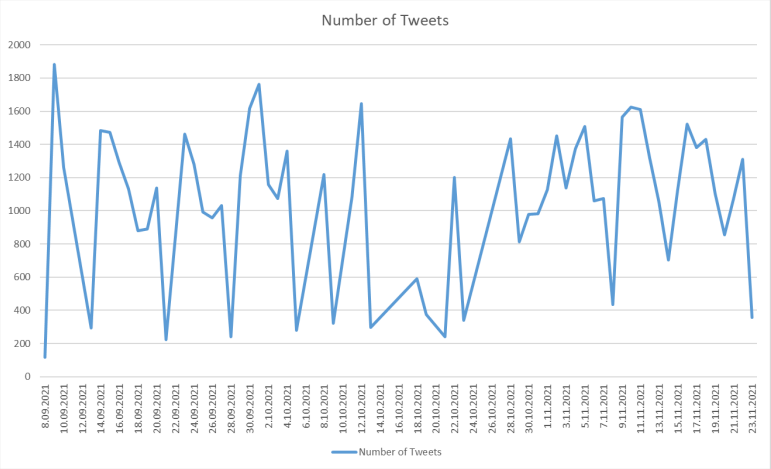

Figure 1. Number of Tweets for the corresponding data range

Users can use a prefix # sign inside the post, which is known as the Twitter hashtag. A hashtag is a tagging standard that makes it easier to find tweets about specific subjects, situations, or events (Ni et al. 2017). The patent related tweet concentration in Hashtags is as shown at table 1. As can be expected, tweets about patents are most intensely gathered under the hashtag #Patent. The second most intense hashtag is #cyrpto (there are also hashtags for nft, bitcoin, blockchain and etc). It will be sufficient to look at the developments in the recent past to explain the crypto technologies being the subject of tweets. Cryptocurrencies first began to function as a means of providing financial transactions between parties. Improvements to crypto technologies over time have enabled real-time and secure transfers, risk mitigation, and enhanced security protocols and regulations. With the rise of these disruptive technologies, there has been a rush to obtain patents for related technologies. This increased the visibility of the cryptocurrency and encouraged new inventors to enter the market and integrate cryptocurrency into new startup ventures. Mentioning Apple's name with patents seems to be closely related to the image created by the brand. Apple

is one of the pioneers of technological innovation. It has launched new and innovative products since 1977, and many companies and business schools around the world have sought to learn about the success story of Apple's innovation (Jun and Sung Park 2013). It is not surprising to see Covid 19 and vaccine brands that are the solution to the desperation it has created on the list.

Table 1. Number of tweets mentioning the given Hashtags

Hashtag	**#of Tweets**	**Hashtag**	**#of Tweets**
#patent(s)	798	#trademark	39
#crypto	97	#verasity	34
#apple	67	#covid19	33
#1	64	#bigpharma	32
#nft	61	#intellectualproperty	31
#moderna	60	#copyright	30
#technology	54	#pfizer	29
#innovation	54	#ps5	28
#bitcoin	49	#cryptocurrency	26
#blockchain	48	#6g	26
#vaccine	48	#xrp	25
#metaverse	46	#pov	24
#5g	44	#invention	24
#news	44	#us	22
#tech	43	#insulin	21
#nfts	43	#pharma	21
#china	39	#btc	21
#tesla	39	#vaccines	20

One of the main purposes of the article is to evaluate the emotional load of the detected tweets. Sentiment analysis has become popular over the last decade, especially for twitter data. In a sentiment analysis, it briefly predicts the sentimental content of texts based on the characteristics defined by the words and expressions used. SentiStrength's concept is based on a list of words that indicates whether a list of words is used in a positive or negative context. Each word in SentiStrength's dictionary is associated with a positive or negative score for the polarity and strength of the emotion term. The score is on a scale of 1 (neutral) to 5 (very positive) or -1 (neutral) to -5 (very negative). For example, on this scale, love gets +3, hate gets -4 points. We have calculated the positive and negative sentiment strength of each matching text and we obtained the following overall average alongside a 95% confidence interval.
Average and 95% confidence intervals:

Pos 1,3751 (1,3698, 1,3804)
Neg 1,3948 (1,3892, 1,4005)
Average positive - Average negative: -0,0197

The strength of the corresponding sentiments is also as illustrated in Figure 2. Very strong sentiments are relatively less when compared to other categories. Strong and very strong negative sentiments are approximately twice more than the positive ones.

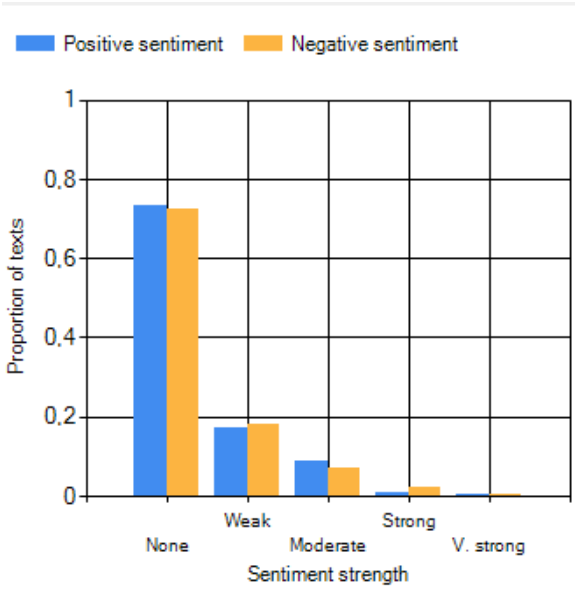

Figure 2. Distribution of the sentiments according to strengths

As can be seen, it is understood that negative emotions are more with a small difference of one percent. When we separate the tweets about the subject according to gender, it is understood that 19.5% of them were posted by males and 9.6% by females. Gender matching could not be done about 70% of the tweets. Because the Mozdeh software extracts gender information from author names, tweets that are gender ambiguous mean that the user's gender cannot be predicted, rather than genderless or third gender. When the emotional intensities were examined by gender, no statistically significant difference was found between the genders.

When the analysis is renewed for the tweets in which vaccines are mentioned together with the patents, it is seen that the emotional loads are collected in the negative direction with a 25% difference. In this case, it is understood that the dominance of negative emotions has increased (1% difference increased to 25%) when it comes to vaccines. Both the decrease in positive emotions and the decrease in negative emotions have a share in the increase in the difference. Interestingly, tweets posted by men are much more negative than women. Table 2 presents the sentiment distribution of tweets on vaccines by country. It is seen that tweets from the United Kingdom are clearly more negative. On the other hand, in tweets about the vaccine, the situation in Australia is quite different from other countries and positive emotions are more intense than negative ones. In the scientific literature, it is possible to come across survey studies conducted in different countries of the world about vaccine hesitancy. In the scientific literature, it is possible to come across survey studies conducted in different countries of the world on vaccine hesitancy(Harapan et al. 2020; Reiter et al. 2020; Soares et al. 2021; Yoda and Katsuyama 2021). In these studies, the rate of individuals who refused to receive the vaccine ranged from 4% to 14.2% (Soares et al. 2021).

Table 2. Country based distribution of average sentiments for vaccine related tweets

Country	**Positive Sentiments**	**Negative Sentiments**	**Difference**
UK	Pos 1,2125 (1,0992, 1,3258)	Neg 1,7000 (1,4968, 1,9032)	Av pos-Av neg:-0,4875
USA	Pos 1,2551 (1,2069, 1,3033)	Neg 1,5823 (1,5124, 1,6521)	Av pos-Av neg:-0,3272
Australia	Pos 1,5667 (1,2944, 1,8389)	Neg 1,4333 (1,1945, 1,6722)	Av pos-Av neg:0,1333
Canada	Pos 1,2706 (1,1513, 1,3899)	Neg 1,5412 (1,3639, 1,7185)	Av pos-Av neg: -0,2706

Table 3 presents the results of the sentiment analysis for tweets focusing on a particular technology or brand from the major hashtags where tweets are collected. Unlike other technologies on the list, Metaverse, crypto and Biontech were mentioned in tweets with a higher positive charge. Metaverse has highest positive load. Metaverse technology is associated with the post-reality universe, a continuous and persistent multi-user environment that combines physical reality with digital virtuality, enabling multi-sensory interactions using virtual environments such as virtual reality

(VR) and augmented reality (AR) and digital objects (Mystakidis 2022). On the other hand, tweets about Moderna are clearly much more associated with negative emotions.

Table 3. Average sentiments for several technologies and brands

Technology	Positive Sentiments	Negative Sentiments	Difference
crypto	Pos 1,3195 (1,2430, 1,3960)	Neg 1,3071 (1,2285, 1,3856)	Av pos-Av neg: 0,0124
bitcoin	Pos 1,2299 (1,1120, 1,3478)	Neg 1,5632 (1,4148, 1,7116)	Av pos-Av neg: -0,3333
Apple	Pos 1,2085 (1,1676, 1,2493)	Neg 1,3254 (1,2764, 1,3744)	Av pos-Av neg: -0,1169
Metaverse	Pos 1,2706 (1,1513, 1,3899)	Neg 1,5412 (1,3639, 1,7185)	Av pos-Av neg: 0,2151
Pfizer	Pos 1,2227 (1,1819, 1,2635)	Neg 1,3393 (1,2881, 1,3906)	Av pos-Av neg: -0,1166
Moderna	Pos 1,2009 (1,1683, 1,2335)	Neg 1,6238 (1,5737, 1,6738)	Av pos-Av neg: -0,4229
Biontech	Pos 1,2273 (1,0115, 1,4430)	Neg 1,0909 (0,9708, 1,2110)	Av pos-Av neg: 0,1364
Johnson& Johnson	Pos 1,3400 (1,1594, 1,5206)	Neg 1,3400 (1,1594, 1,5206)	Av pos-Av neg: 0,0000

When the patent-themed Tweets discussed in this study are analyzed according to users, it is understood that the person who tweets the most is Elon MUSK with 382 tweets. In second place, an American politician and activist named Bernard Sanders with 177 Tweets.

4. Conclusion

Patents are official documents that provide information about the direction and pace of technological development. Although the analysis of these documents is frequently encountered in the literature, the issue of how these documents is perceived by the public has not yet been addressed. This study can be seen as the first demonstration of an initial effort in this regard. The study is based on the analysis of tweets that the society commonly prefers to communicate. The study showed that patents are topics that attract the attention of the public, but in particular, patents for some technologies seem to be more popular than others. Cryptosystems have often been the subject of tweets about patents. The sentimental charge of these tweets is generally positive. Similarly, vaccine patents were also frequently discussed in tweets. However, the sentimental load in vaccine patents is not as positive as in crypto systems. It essential to state that this research was conducted with limited data. It is not possible to generalize the results as it is not possible to represent the entire population. It is possible to improve this work with a deeper analysis and a more comprehensive data set.

References

Ahmed, W., Bath, P. A., Sbaffi, L., & Demartini, G. (2018). Measuring the Effect of Public Health Campaigns on Twitter: The Case of World Autism Awareness Day. In G. Chowdhury, J. McLeod, V. Gillet, & P. Willett (Eds.), *Transforming Digital Worlds* (pp. 10–16). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-78105-1_2

Bibi, M., Abbasi, W. A., Aziz, W., Khalil, S., Uddin, M., Iwendi, C., & Gadekallu, T. R. (2022). A novel unsupervised ensemble framework using concept-based linguistic methods and machine learning for twitter sentiment analysis. *Pattern Recognition Letters*, *158*, 80–86. https://doi.org/10.1016/j.patrec.2022.04.004

Crannell, W. C., Clark, E., Jones, C., James, T. A., & Moore, J. (2016). A pattern-matched Twitter analysis of US cancer-patient sentiments. *Journal of Surgical Research*, *206*(2), 536–542. https://doi.org/10.1016/j.jss.2016.06.050

Durmuşoğlu, A. (2017). Effects of Clean Air Act on Patenting Activities in Chemical Industry: Learning from Past Experiences. *Sustainability*, *9*(5), 862. https://doi.org/10.3390/su9050862

Durmusoglu, A., & Unutmaz Durmusoglu, Z. D. (2021). Traffic Control System Technologies for Road Vehicles: A Patent Analysis. *IEEE Intelligent Transportation Systems Magazine*, *13*(1), 31–41. Presented at the IEEE Intelligent Transportation Systems Magazine. https://doi.org/10.1109/MITS.2020.3037319

Giachanou, A., & Crestani, F. (2016). Like It or Not: A Survey of Twitter Sentiment Analysis Methods. *ACM Computing Surveys*, *49*(2), 28:1-28:41. https://doi.org/10.1145/2938640

Harapan, H., Wagner, A. L., Yufika, A., Winardi, W., Anwar, S., Gan, A. K., et al. (2020). Acceptance of a COVID-19 Vaccine in Southeast Asia: A Cross-Sectional Study in Indonesia. *Frontiers in Public Health*, *8*. https://www.frontiersin.org/article/10.3389/fpubh.2020.00381. Accessed 28 April 2022

Jun, S., & Sung Park, S. (2013). Examining technological innovation of Apple using patent analysis. *Industrial Management & Data Systems*, *113*(6), 890–907. https://doi.org/10.1108/IMDS-01-2013-0032

Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology & Politics*, *13*(1), 72–91. https://doi.org/10.1080/19331681.2015.1132401

Mystakidis, S. (2022). Metaverse. *Encyclopedia*, *2*(1), 486–497. https://doi.org/10.3390/encyclopedia2010031

Ni, M., He, Q., & Gao, J. (2017). Forecasting the Subway Passenger Flow Under Event Occurrences With Social Media. *IEEE Transactions on Intelligent Transportation Systems*, *18*(6), 1623–1632. Presented at the IEEE Transactions on Intelligent Transportation Systems. https://doi.org/10.1109/TITS.2016.2611644

Paul, M. J., Sarker, A., Brownstein, J. S., Nikfarjam, A., Scotch, M., Smith, K. L., & Gonzalez, G. (2015). Social media mining for public health monitoring and surveillance. In *Biocomputing 2016* (pp. 468–479). WORLD SCIENTIFIC. https://doi.org/10.1142/9789814749411_0043

Reiter, P. L., Pennell, M. L., & Katz, M. L. (2020). Acceptability of a COVID-19 vaccine among adults in the United States: How many people would get vaccinated? *Vaccine*, *38*(42), 6500–6507. https://doi.org/10.1016/j.vaccine.2020.08.043

Soares, P., Rocha, J. V., Moniz, M., Gama, A., Laires, P. A., Pedro, A. R., et al. (2021). Factors Associated with COVID-19 Vaccine Hesitancy. *Vaccines*, *9*(3), 300. https://doi.org/10.3390/vaccines9030300

Tumasjan, A., Braun, R., & Stolz, B. (2021). Twitter sentiment as a weak signal in venture capital financing. *Journal of Business Venturing*, *36*(2), 106062. https://doi.org/10.1016/j.jbusvent.2020.106062

Wei, Y., Gong, P., Zhang, J., & Wang, L. (2021). Exploring public opinions on climate change policy in "Big Data Era"—A case study of the European Union Emission Trading System (EU-ETS) based on Twitter. *Energy Policy*, *158*, 112559. https://doi.org/10.1016/j.enpol.2021.112559

Yoda, T., & Katsuyama, H. (2021). Willingness to Receive COVID-19 Vaccination in Japan. *Vaccines*, *9*(1), 48. https://doi.org/10.3390/vaccines9010048