Investigating Public Perception of ChatGPT: A Twitter Sentiment Analysis

Hadeel Fahad Altowairqi
Department of Computer Science, Exeter University
Ha562@exeter.ac.uk

I. INTRODUCTION

Since it was released, ChatGPT has flooded the internet and caused public controversy involving people from all parts of society. This large language model (LLM) created by Open AI functions as an interactive chatbot that generates text for a variety of tasks such as academic writing, language translations, and coding. [1] With its impressive ability to produce a comprehensible human-like text, ChatGPT has become a popular tool for both researchers and public users. Therefore, Understanding the social context surrounding technology is crucial for future adoption and use.

Several studies have been conducted to investigate the public's perception of chatbots to understand their potential benefits in different fields. For example, in marketing, knowledge of people's attitudes toward chatbots in mobile advertising can help businesses develop and test new strategies. Similarly, the way teachers view chatbots can influence their willingness to integrate this technology into their work. [2]

Social media platforms, particularly Twitter, offer a rich source of data that can be utilized to examine public attitudes towards ChatGPT. Twitter allows individuals to post tweets, which are 280-character messages expressing their thoughts and opinions. [3] This serves as a valuable data source for sentiment analysis, a natural language processing (NLP) technique that helps determine the sentiment or emotional tone underly the text. This involves categorizing text as positive, negative, or neutral based on the sentiment conveyed by the words used. [4]

II. PROBLEM STATEMENT

This project extends past research on general sentiment towards ChatGPT by examining sentiment around ChatGPT and getting insights on tweet topics and user occupations as well. This analysis will help to identify gaps in knowledge and misinformation surrounding large language models (LLMs) like ChatGPT. which can be used to improve the educational efforts and communication strategies to better inform the public. Furthermore, project findings can help formulate future directions for the research and development of ChatGPT and other LLMs. Ultimately, public opinion plays a crucial role in the integration of these models into various sectors, making a nuanced understanding of user sentiment essential. In particular, I asked the following research questions:

RQ1. What is the general sentiment towards Chat GPT?

RQ2. What are the main topics that are being discussed about ChatGPT on Twitter?

RQ3. What are the characteristics of ChatGPT early users in terms of user occupation?

Initially, I will use a sentiment analysis model on tweets related to ChatGPT to get the general sentiment. Following this, I will employ a topic modeling method to identify the most frequently discussed topics. Subsequently, I will categorize the tweets by occupation using the user descriptions to get more detailed insights.

III. RELATED WORK

Sentiment analysis within the domain of micro-blogging could be a relatively new research topic. Several studies have focused on sentiment analysis of social media data, including Twitter data from different viewpoints. [5] discussed the broad social impact of ChatGPT using social science methods such as interviews and user-experience analyses. Similarly [6][7] provide an expert view of the domain-specific impact of using ChatGPT in engineering and healthcare.

In terms of computational analysis, [8] used Latent Dirichlet Allocation (LDA) to identify popular topics in ChatGPT-related tweets, then performed sentiment analysis on a small dataset by manually labeling the sentiment on the identified topics. More recently [1] applied sentiment analysis to investigate public opinion, and classified tweets into 19 predetermined topics with a **roBERTa**-based model fine-tuned for tweet topic classification. However, they did not include demographic analysis on the dataset.

Here, I aim to study the general attitude towards ChatGPT using Twitter data by performing sentiment analysis as well as discovering the major topics discussed around ChatGPT with topic modeling techniques. Then use the demographic information to understand how occupations influence the sentiments towards ChatGPT.

IV. METHODOLOGY

This section describes the methodology in detail, starting from acquiring the dataset to analysis methods. first, the data was cleaned by removing irrelevant data such as URL links and emojis. Then, the tweets were preprocessed for sentiment analysis and topic modeling. The overview of the workflow is shown in **Figure 1.**

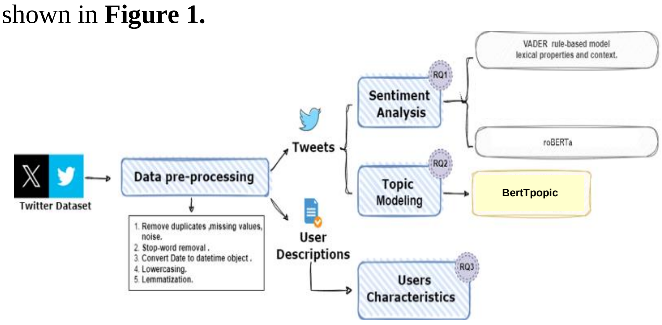


Figure 1 General workflow.

a. Dataset

The dataset used in this project is publicly available in Kaggle. it contains tweets collected with the hashtag (#Chatgpt) for the period between November 30, 2022, and Aprile 8,2023. The dataset has 478,347 tweets along with their corresponding username, user description, user location, and timestamp information.

b. Data Cleaning

To prepare data for sentiment analysis, and topic modeling the following preprocessing steps are applied to clean and prepare the dataset:

- Breaking down text into smaller pieces called tokens using Tokenization.
- Lemmatization of the tweets to get the base or lemma.
- Removing links, mentions, stopwords, and non-words.
- Lowercasing the text.
- Tokenization. And Lemmatization

c. Sentiment Analysis

To answer the first research question regarding general sentiment, two models were chosen. The first one is **VADER**, a rule-based sentiment analysis tool specifically designed for analyzing social media texts. It is a pre-trained sentiment analysis model that uses a dictionary of words and rules to determine the sentiment of a piece of text [9]. The second model is **RoBERTa** which is an optimized version of the **BERT** transformer developed by Cardiff University available on huggingface. **RoBERTa** analyses text by considering the context of words in both directions (left and right of each word) [10].In this project, I used a pre-trained version of **RoBERTa**. This approach is useful when we have a large volume of data and want to leverage pre-trained models to avoid the need for extensive feature engineering.

d. Topic Modeling

To answer the second research question, I used the **BERTopic** technique which is a topic modeling approach that utilizes transformers and c-TF-IDF to generate dense clusters. **BERTopic** used these clusters to identify topics while keeping important words in the topic descriptions.[11]

e. Occupation Extraction and Analysis

To extract the user occupation from user descriptions I used unigrams and bigrams of the preprocessed texts to match the occupation list. The list of occupations obtained from the Standard Occupational Classification (SOC) contains more than 23 occupation descriptions from different fields.

V. Results

a. Sentiment Analysis

The majority of tweets, according to the sentiment analysis, are positive about ChatGPT and varied over time as shown in **Figure 4**. Users frequently emphasize how well it can produce responses that are both logical and appropriate for the given context. A small percentage of tweets, nevertheless, raised issues about possible technological abuse and ethical considerations.

VADER identified nearly half (49.8%) of tweets as positive, highlighting user satisfaction with ChatGPT's logical and relevant responses. However, a small portion (11.6%) expressed concerns, likely about potential misuse. However, Roberta (used on a small random sample due to limitations), showed similar trends in **Figure 2** with positive sentiment (29.9%) but also a higher percentage of negative tweets (20.2%). While both models suggest a generally positive view, Roberta might be more sensitive to concerns in the data.

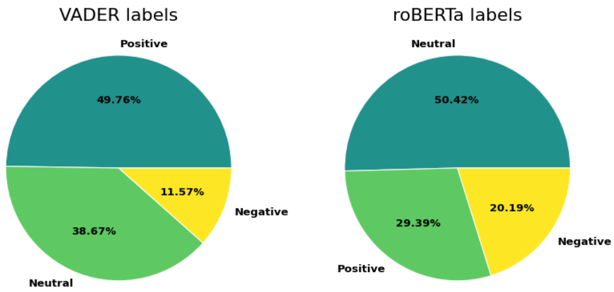


Figure 2 VADER and Roberta model Sentiment Results.

Figure 3 compares the polarity score between the two models and it highlights that VADER model tends to classify most tweets as positive. Whereas, Roberta presents more balanced results as it trained on similar data sets which help it to extract the underlying meaning around the words.

b. Topic Modeling

The **BERTopic** model extracts 952 topics 7 of them related to ChatGPT. Due to the limited computational resources, I applied topic analysis on a random sample which counts about (50%) of the dataset. The tweet count for topics ranges from 10 to 43,657. The topic with the highest tweet count is ‘Artificial Intelligence’. **Table 1** displays the tweet count and example for each top topic.

Topic	Count	Example
Artificial intelligence	43657	OpenAI develop ChatGPT artificial intelligence that changes everything
language model	1569	AI language model capable of generating human-like response on a wide range of topic
Google Bard	1313	Google launched a competitor to ChatGPT, it looks like ChatGPT
Education	761	recent survey shows how ChatGPT is used by teachers and students in teaching and learning
Job replacement	579	ai going take jobs, ChatGPT replace job
Writing	327	use ChatGPT as a writer, AI technology offers writers a better way to write
Healthcare	258	Researcher intends to use ChatGPT to improve healthcare and provide accurate medical information

Table 1 Top 7 topics with word count and examples

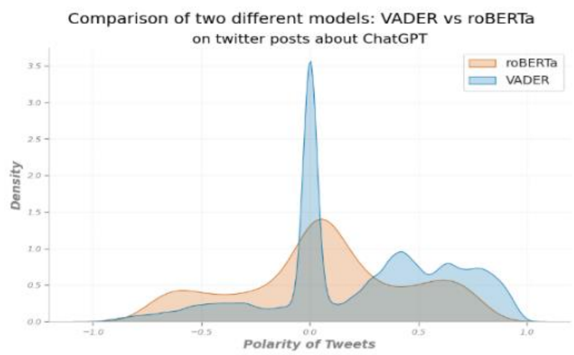


Figure 3 Polarity score Distribution.

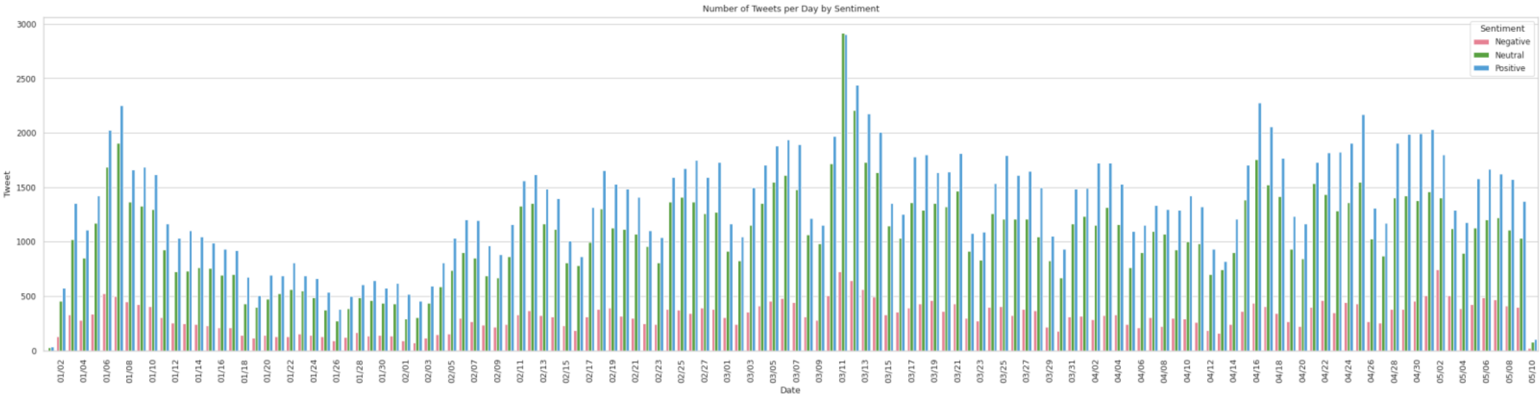


Figure 4 sentiment over time

c. Occupation Extraction

About (184,338) occupation titles were extracted from the dataset. (71.8%) of them either did not match the occupation list or failed to extract the title. Occupations are aggregated

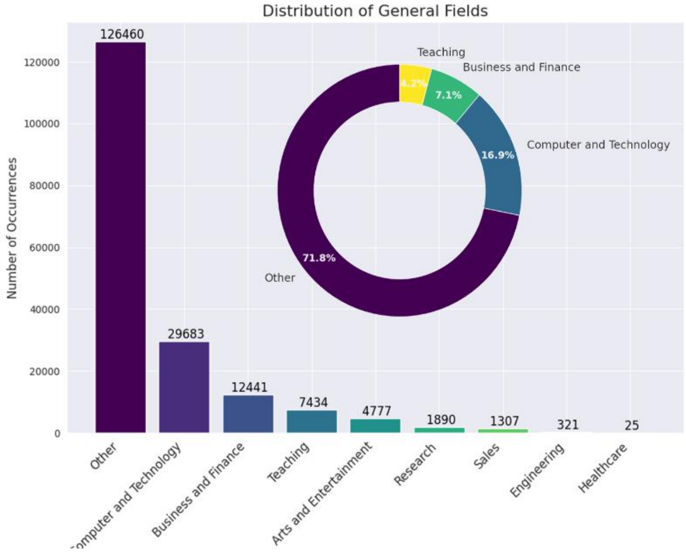


Figure 5 Distribution of occupations.

into larger fields to analyze the overall spread of job types within the dataset. **Figure 5** visualizes the distribution of professions among those tweeting about ChatGPT. Users in computer and technology make up the largest group (16.9%), followed by business and finance (7.1%). Healthcare professionals seem to be the least engaged, less than (1%). A detailed representation of occupation types and counts can be rich in the Appendix.

VI. DISCUSSION

The overall sentiment result has a higher percentage of positive and neutral tweets compared to the work by Leiter et al. 2023[1], which has 29.9% positive, 52.2% neutral, and 17.9% negative tweets. This difference could be due to several factors, including sampling methods and data cleaning techniques. The use of Roberta, a powerful sentiment analysis model, might have also played a role. However, it's important to consider that my analysis might not be directly comparable. Further investigation with identical datasets and models would provide a more definitive picture.

Using BERTopic, I identify the top topics related to ChatGPT on Twitter, as shown in **Table 1**. The top topics can be grouped into three main categories. The first group focuses on the future of AI (Artificial Intelligence and 'Language Models'**)**. The arrival of ChatGPT has ignited public discourse about artificial intelligence and its impact. Tweets in this category discuss how ChatGPT and advancements in AI will shape our lives, exploring both the technological aspects and ChatGPT's broader implications. The second group is the topics around applications of ChatGPT. The topics in this group are 'Writing' and 'Healthcare. Tweets in this group are basically about how people can leverage ChatGPT to either answer basic day-to-day questions or solve complex problems in their domains. The third category delves into potential disruptions caused by ChatGPT including topics about "Education," "Job Replacement," and "Google Bard". Discussions within "Google Bard" center on whether ChatGPT could replace search engines or become an integrated part of them. Similarly, tweets under "Education" and "Job Replacement" explore the potential of ChatGPT as a learning tool and its influence on future job landscapes.

The occupation analysis reveals that the top occupation group posted about ChatGPT is 'computer and technology '.As expected when we have new technology people from the same field start to investigate this technology. Art and entertainment workers and teachers started to be curious about the applications of ChatGPT in art and teaching, which suggested new fields of applications and improvements. However, in health care ChatGPT and other LLMs at that time did not have any applications due to their sensitivity. This reflects the small portion of tweets from healthcare workers. and suggest the potential opportunity to integrate ChatGPT in this field.

VII. CONCLUSION

In conclusion, this exploratory investigation of Twitter discourse surrounding ChatGPT gives multifaceted insights. Sentiment analysis revealed a predominantly positive effect, with users commending ChatGPT's ability to generate logically coherent and contextually relevant responses. While some concerns exist regarding potential misuse and ethical considerations Additionally, topic modeling identified various potential applications for ChatGPT, spanning from creative writing and healthcare to education and business. These findings highlight the versatility of ChatGPT and its potential future impact on various sectors. However, some discussions surrounding potential disruptions caused by ChatGPT, including job displacement and educational integration, necessitate further exploration. **Limitations:** It's important to acknowledge the limitations of this study. The use of a random sample for topic modeling and potential biases in sentiment analysis models need further investigation with a larger, unbiased dataset.

REFERENCES

[1] Christoph Leiter, Ran Zhang, Yanran Chen, Jonas Blouadi, Daniil Larionov, Vivian Fresen, and Steffen Eger. 2023. Chatgpt: A meta-analysis after 2.5 months. arXiv preprint arXiv:2302.13795.

[2] Lucrezia Maria de Cosmo, Luigi Piper, and Arianna Di Vittorio. 2021. The role of attitude toward chatbots and privacy concerns on the relationship between attitude toward mobile advertising and behavioral intent to use chatbots. Italian Journal of Marketing, 2021:83–102.

[3] Wikipedia contributors. (2024, March 13). *Twitter*. Wikipedia. https://en.wikipedia.org/wiki/Twitter

[4] Kharde, V. A., & Sonawane, S. (2016). Sentiment Analysis of Twitter Data: A Survey of Techniques. *International Journal of Computer Applications*, *139*(11), 5–15. https://doi.org/10.5120/ijca2016908625

[5] Malak Abdullah, Alia Madain, and Yaser Jararweh. 2022. Chatgpt: Fundamentals, applications, and social impacts. In 2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS), pages 1–8. IEEE.

[6] Som SBiswas. 2023. Role of chat gpt in public health. Annals of Biomedical Engineering, pages 1–2.

[7] Yogesh K Dwivedi, Nir Kshetri, Laurie Hughes, Emma Louise Slade, Anand Jeyaraj, Arpan Kumar Kar, Abdullah M Baabdullah, Alex Koohang, Vishnupriya Raghavan, Manju Ahuja, et al. 2023. "so what if chatgpt wrote it?" multidisciplinary perspectives on opportunities, challenges and implications of generative conversational ai for research, practice and policy. International Journal of Information Management, 71:102642.

[8] Mubin Ul Haque, Isuru Dharmadasa, Zarrin Tasnim Sworna, Roshan Namal Rajapakse, and Hussain Ahmad. 2022. " I think this is the most disruptive technology": Exploring sentiments of chatgpt early adopters using Twitter data. arXiv preprint arXiv:2212.05856.

[9] Hutto, C. J., & Gilbert, É. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

[10] Francesco Barbieri, Jose Camacho-Collados, Leonardo Neves, and Luis Espinosa-Anke. 2020. Tweet eval: Unified benchmark and comparative evaluation for tweet classification. arXiv preprint arXiv:2010.12421.

[11] Grootendorst, M. (2022, March 11). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv.org. https://arxiv.org/abs/2203.05794.

Appendix

Exploratory Data Analysis (EDA):

Locations by Total Tweets :

The following figure shows the top 10 locations in terms of number of tweets. As expected, the USA produced the highest number of tweets more than (30000 tweets). The United Kingdom and India have similar numbers near 150000 tweets. Whereas many tweets more than 15K did not have location information.

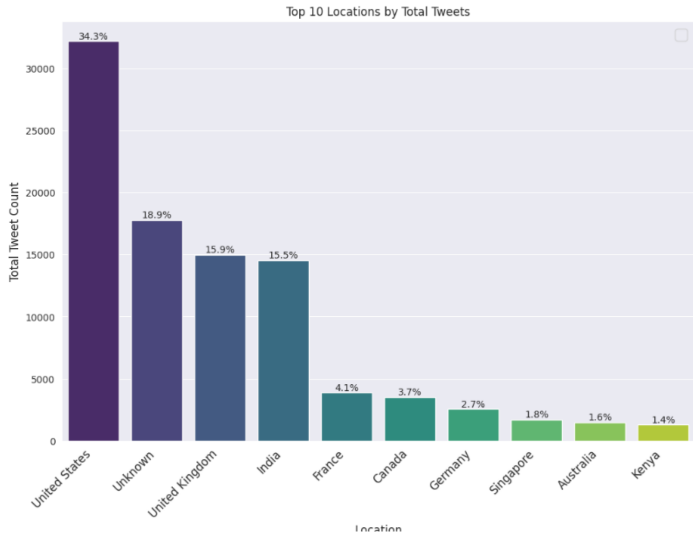


Figure 2

Timeline counts of tweets:

The following figure shows the timeline distribution From 30 November to 4 April, we can observe several packs which match the days when a significant event related to Ai was released or new updates and features were added to chatgpt. For example, in February 2023 Google launched Bard the competitor of ChatGPT, and when

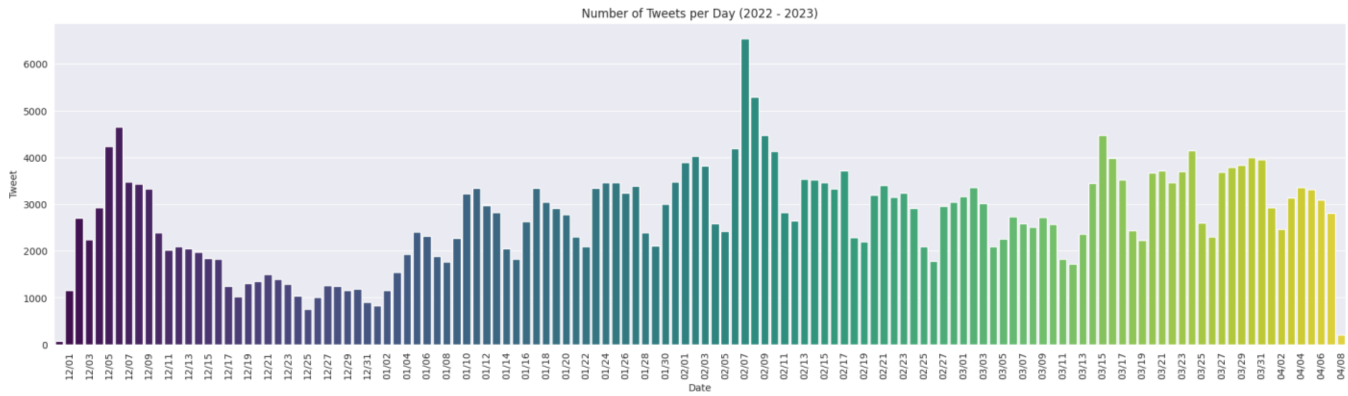
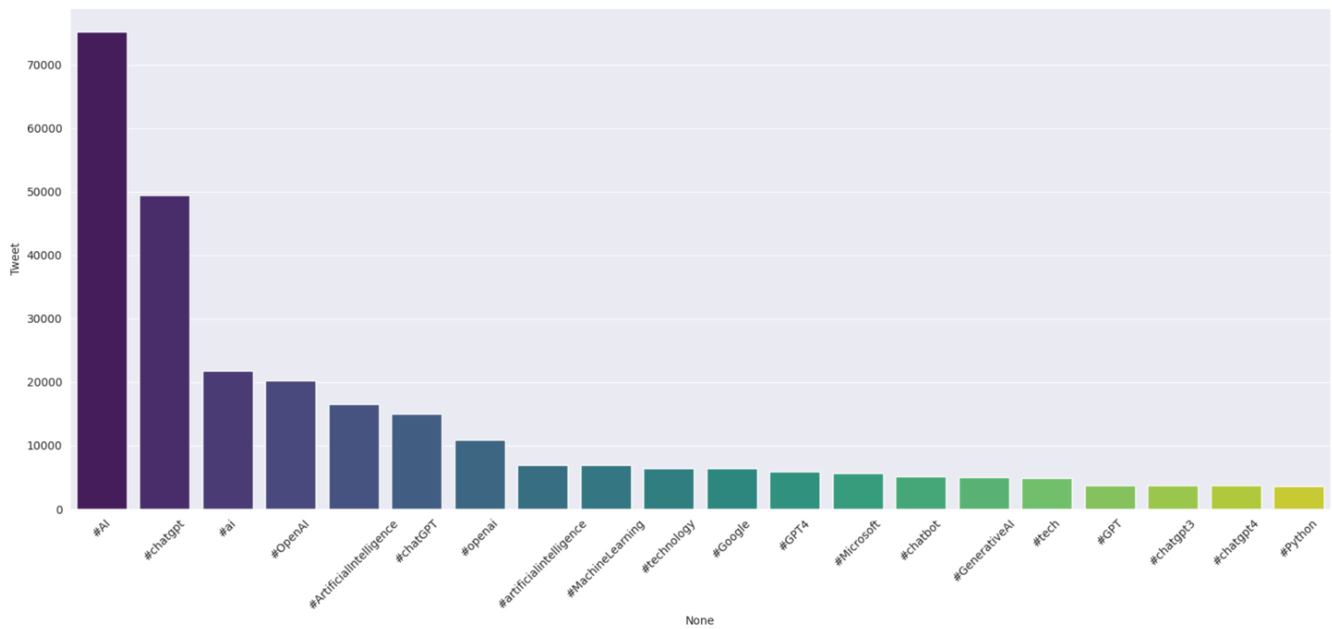


Figure 3 tweets timeline.

ChatGPT reached 100 million monthly active users in January, just two months after its launch.

Top hashtags used in tweets about ChatGPT



Top 20 occupations extracted from user description.

People from different fields tweets about ChatGPT, most of them related to technology and engineering as shown in **Figure 6** .

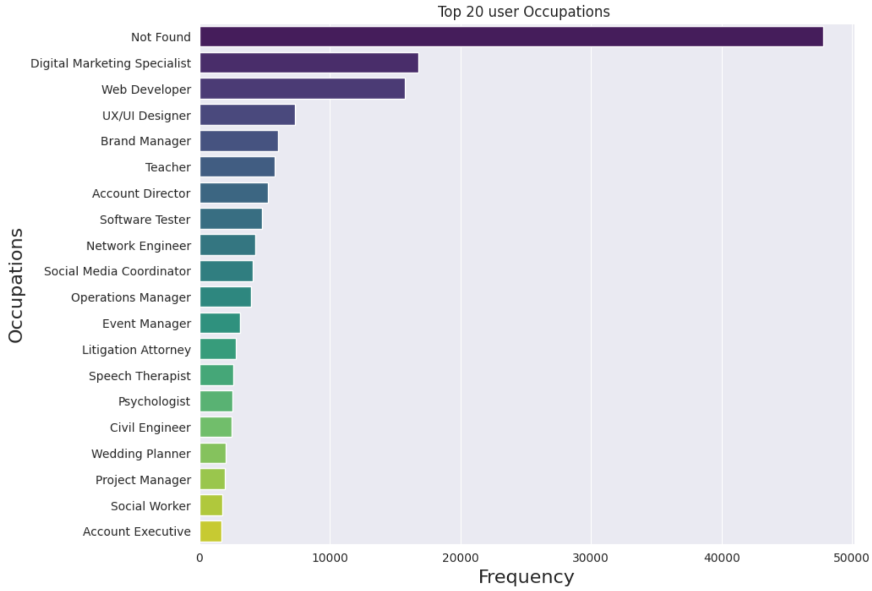


Figure 4 user occupations.

Top Topics with word representation :

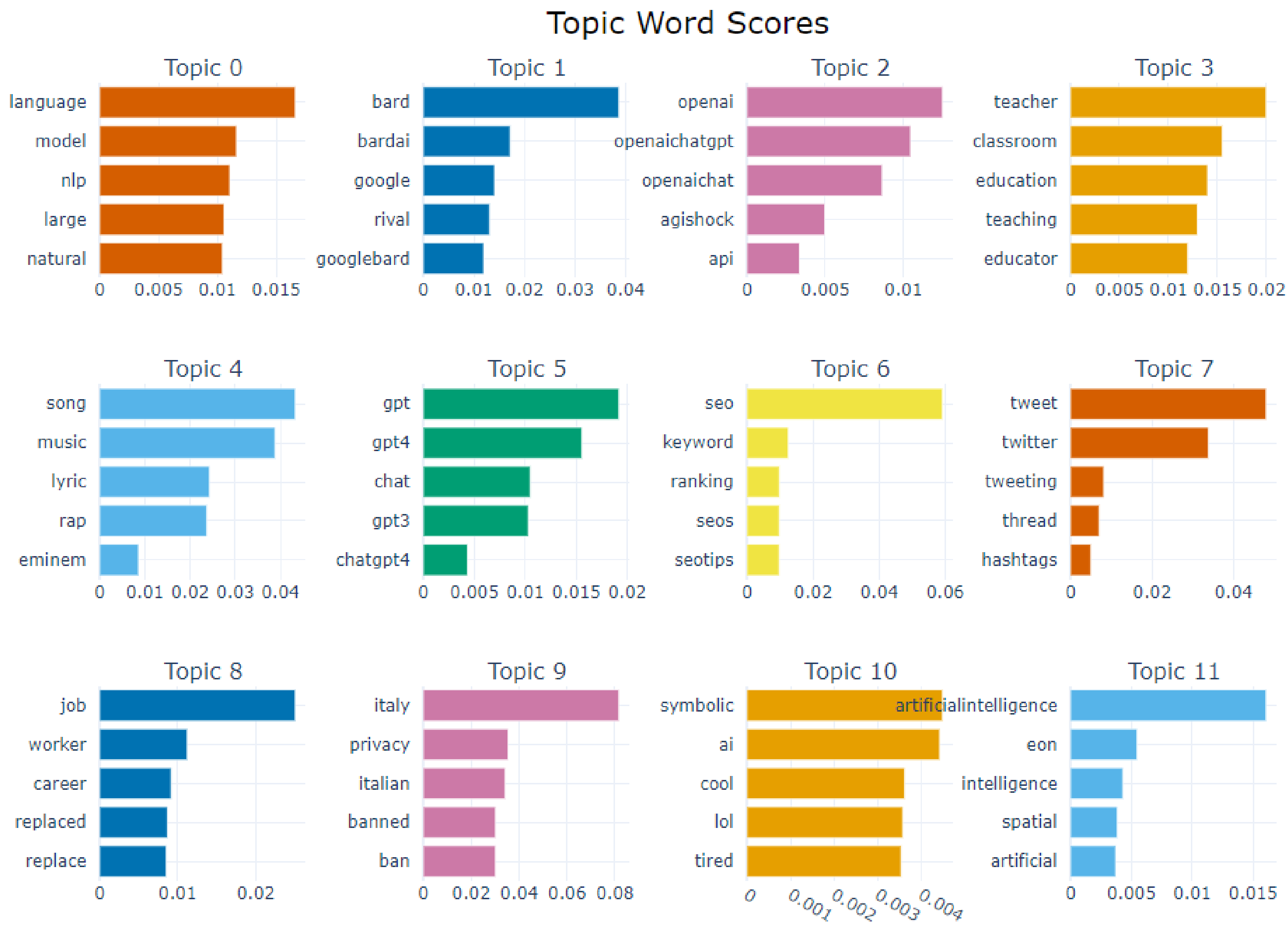


Figure 5 Top 12 topics extracted from the dataset with their representative words