Dec 11th, 12:00 AM

An Investigation of Domain-based Social Influence on ChatGPT-Related Twitter Data

Supunmali Ahangama
University of Moratuwa, supunmali@uom.lk

An Investigation of Domain-based Social Influence on ChatGPT-Related Twitter Data

Completed Research Paper

Supunmali Ahangama
Department of Information Technology, Faculty of Information Technology,
University of Moratuwa
Bandaranayake Mawatha, Moratuwa 10400, Sri Lanka
supunmali@uom.lk

Abstract

Recently, the word "ChatGPT" has made its way into the vocabulary of various fields that are important to society, including healthcare, education, governance, and robotics. This study examines the various reasons for social media influences using the Social Influence Theory. The study used tweets connected to ChatGPT collected from January 1st, 2023 to March 31st, 2023, totaling 402, 965 tweets. Initially, the key domains of ChatGPT discussions were identified along with the variations in the sentiments over the period. Subsequently drawing the literature related to Social Influence Theory, this study explored the relationship between subjective norm, social identity, user commitment, and features of the tweet content with social influence on the readership of a ChatGPT-related tweet. Subjective norm, identity, sentiment and domain are significant factors for social influence. The user commitment suggests a negative relationship with the social influence. Finally, the theoretical and practical implications are explained.

Keywords: ChatGPT, social influence, social media, identity, Twitter, topic modelling

Introduction

Advanced large language models (LLMs) like Generative Pre-trained Transformer (GPT-3) are presently trending in an unprecedented manner (Kasneci et al., 2023) and being applied in many different fields spanning from computer science to public health, marketing and education as state-of-the-art technology (Androutsopoulou et al., 2019; Biswas, 2023a, 2023b; Haque et al., 2022; Taecharungroj, 2023; Tinholt, D., 2017). To learn the human language and respond to human-like text or perform other natural language-related activities, these LLMs are trained on very large datasets (e.g. Wikipedia articles, web pages, and books) and fine-tuned to specific activities (Dwivedi et al., 2023). The latest version of Chat Generative Pre-trained Transformer (ChatGPT), a chatbot launched by OpenAI in November 2022, had reached over 100 million users by the end of January 2023 (Hu, 2023). Presently, GPT-3 and GPT-4 text interpreters are used in ChatGPT. Compared to earlier chatbots, ChatGPT is capable of generating natural language responses to user queries while personalizing based on the previous conversation (Kasneci et al., 2023). In addition, numerous experiments have been run to confirm the quality of ChatGPT's outputs. As an illustration, several news articles published that ChatGPT was able to complete graduate-level exams related to law and business (Kelly, 2023) successfully and was even placed among the top 10% of students (Koetsier, 2023).

Due to the substantial growth in popularity and different use cases, many have started discussing and sharing their experiences of using ChatGPT for various tasks like generating answers for questions, creative work like poems and stories, and having conversations on social media. Thus, considering this hype, a lot of discussions and knowledge dissemination are taking place successfully on social media and microblogging platforms. Social media is a useful tool for gathering public opinion quickly because of its inherent properties that allow information to be quickly disseminated (Haque et al., 2022). Social media sites like Twitter are rich in information related to ChatGPT as they allow (a) gathering an exhaustive range of viewpoints and feedback from an extensive array of people from a wide range of backgrounds and perspectives rather than relying solely on a selected group of experts (target audience); (b) extracting useful information to understand the influence and reach of ChatGPT among the users; and (c) assessing the public opinions relevant to a topic or an event quickly (Korkmaz et al., 2023). Additionally, it was decided to base the study on Twitter data because many individuals can concurrently give their opinions on current affairs that are now trending both locally and internationally via Twitter (Korkmaz et al., 2023). These reasons influenced the extraction of user-generated content related to ChatGPT from Twitter to explore the relationships among users and the text contents (including sentiment) of tweets.

There is a dearth of research related to ChatGPT. Going through the available literature, it can be observed that there is a gradual increase in studies carried out related to GPT-3. However, there are several research gaps. First, many of these studies focus on the technical aspects of it and are used for comparison of various AI chatbot techniques and deep learning models (Biswas, 2023a; Howard et al., 2023; Patel & Lam, 2023; Surameery & Shakor, 2023). Though such studies are very useful for the continuous improvements in the LLMs, it is also important to understand and analyse the user perceptions, application areas and use cases of such systems. For the generalizability of these models, it is crucial to explore and investigate the use cases and user perceptions (Taecharungroj, 2023). Thus, considering the limited studies available, it was decided to investigate the user perceptions and applicability of them in this study. There are several studies carried out regarding its applications in health (Biswas, 2023a; Patel & Lam, 2023; Qi et al., 2023), education (Farrokhnia et al., 2023; Kasneci et al., 2023; Tlili et al., 2023), and climate change (Biswas, 2023b). Second, several studies related to social media have explored mainly the sentiment (Korkmaz et al., 2023) and topics being discussed (Li et al., 2023) rather than investigating the social influence of those tweets. Third, those studies have not been done considering a large dataset of tweets related to ChatGPT or they have only considered tweets related to a specific domain like education (Li et al., 2023).

Even though there are large amounts of data, many of which are not that useful to the user. Thus, it would be useful to identify the most influential tweets to the users. There are several benefits of exploring the social influence of tweets on users. First, understanding the perspectives of early adopters of new technology (like ChatGPT) would be helpful as their shared thoughts can influence how new technology is perceived more widely. Second, it would be useful to understand the key domains in which these technologies can be applied and their strengths and weaknesses. Finally, identifying the key factors that could make a tweet to be influential would be assistive in understanding how information and ideas propagate in social media. It was decided to use the Social Influence Theory (Kelman, 1958) to explore the parameters for popularizing and using ChatGPT. Here, social influence refers to how people can influence others to change their ideas and behaviour.

However, ICT can be considered a vehicle for creating mobilizing structures (Earl and Kimport 2011). Here, the role of social media (namely, Twitter) is considered to understand the influence of individual online activities in generating a society with better societal participation. It would be useful to evaluate user perceptions and how they are influential on other users. Several studies have evaluated the social influence based on Twitter content, users and networks (Oldeweme et al., 2021; Venkatesan et al., 2021). Furthermore, some of the studies have explored the factors influencing word of mouth in online communities (Bakshy et al., 2011; Barbagallo et al., 2012). Moreover, it has been indicated that influence is not the same for all the topics and there are different influences on different domains (Alizadeh et al., 2020; Xiao et al., 2014). Nonetheless, there have been fewer studies focusing on the influence within the specific topic (or domain) (Hamzehei et al., 2017). Thus, it is planned to explore how Social Influence Theory can be applied in the context of the introduction of ChatGPT in this study.

In this paper, the following research questions (RQs) will be addressed.

RQ 1: What are the topics being discussed on Twitter by the early adopters of ChatGPT?

RQ2: What are the factors that contribute to creating social influence based on the content related to ChatGPT on Twitter

In this study, 402, 965 tweets in English within the period from January 1st 2023 to 31st March 2023 were extracted. Initially, a word cloud was prepared using the dataset to investigate RQ 1. Moreover, topic modelling was done using BERTopic to identify the common themes (or topics) in tweets. Also, a sentiment analysis was performed using Valence Aware Dictionary and sEntiment Reasoner (VADER). lexicon and rule-based sentiment analysis tool. Subsequently, to explore RQ2, a set of hypotheses would be developed and analysed for social influences of tweets on ChatGPT. Here, the parameters relevant to the user (availability of verified account, follower count, etc.) and content (sentiment, domains) of the tweet were used.

Important contributions of the paper can be identified as follows. It provides an analysis of the topics being discussed on Twitter related to ChatGPT. The findings could be generalized into any other state-of-the-art technology being introduced and will be useful in understanding the perceptions of the early adopters. Usually, the early adopters pave the way for future growth as well as the direction of an IT artefact (Haque et al., 2022; Taecharungroj, 2023). Further, a domain-based study would be undertaken to determine how the various elements of user profiles, as well as the content being shared, can lead to increased social influence. The amount of views on a post represents its social influence. It is possible to determine whether content sentiments, as well as the number of followers and the length of time the account has been active, will be influential. Finally, the use of theory related to Social Influence would be useful in extending the theoretical discourse related to social media usage in the presence of social influence.

In the next section, the limited studies related to ChatGPT will be discussed along with the applicability of the Theory of Social Influence in the social media context. In the next section, the research model will be explained with the hypotheses. Subsequently, the dataset and the methodology followed will be elaborated. Next, the analysis of the results followed by the discussion on implications and limitations will be given along with the Conclusion.

Theory and Hypothesis Development

Considering the disruptive nature, many studies are being done regarding the impact and application of ChatGPT in various domains. Table 1 summarises several key publications done after the public release of ChatGPT. Even though there are many studies, all of which have considered the evaluation of response generated by ChatGPT for the given queries. There are only a very few publications to evaluate the user perception related to this technology.

Domain	Areas Explored
Education	Impact of critical thinking, data privacy, plagiarism and assessment design for higher education (Kasneci et al., 2023; King, 2023; Li et al., 2023; Tlili et al., 2023). Personalized and interactive learning and reducing teaching workload (Farrokhnia et al., 2023; Kohnke et al., 2023)
Health	Creation of discharge summaries for patients in hospitals (Patel & Lam, 2023) Challenges in using in clinical practice (Howard et al., 2023) Making informed decisions about individual health and use cases in public health (Biswas, 2023a) Applicability in nursing education (Qi et al., 2023) Review of articles published in PubMed and Google Scholar related to healthcare (Sallam, 2023).
Computing	Application in debugging and bug fixing (Surameery & Shakor, 2023) Creating and modifying software architectures (Ahmad et al., 2023)
Cybersecurity and Privacy	Threats of using ChatGPT (Gupta et al., 2023) and responses generated from ChatGPT to face the challenges (Mijwil et al., 2023)
Environment	Use cases for handling global warming based on responses received from ChatGPT (Biswas, 2023b) Provision of incomplete safety advice (Oviedo-Trespalacios et al., 2023)
Academia and Research	Impact on the generation of scholarly knowledge and libraries based on responses received from ChatGPT (Liebrenz et al., 2023; Lund & Wang, 2023).

Table 1. Research Done in Different Domains

Social media and micro-blogging systems like Twitter have created massive amounts of data and it has revolutionised the way information is used and disseminated. Three main components can be identified in a social media platform as a socio-technical system (Cho & Wash, 2021). First, the technologies used such as the interactive web/mobile-based system facilitating communication among the members, content sharing and archival and user profiles can be considered. The second component comprises the people who participate and interact in the community. Finally, the content available for others to consume can be identified. On Twitter, most of the time this content is available for the public to read. Thus, it would impact the perception of users of certain cutting-edge technologies. It is possible to identify the domain of knowledge-based expertise and experience in various areas like sports, computer science, education, entertainment, etc. However, in the social media context, the domain can be identified based on the type of content generated by the general public. There have been few studies done to understand the topics being discussed in social media (Taecharungroj, 2023). The Latent Dirichlet Allocation (LDA) model has been used in the study (Taecharungroj, 2023) and they have identified news, technology, and reactions as the key topics discussed from 30th November to 31st December 2022. Thus, it would be useful to evaluate the interests of the general public related to ChatGPT. Furthermore, there is a minuscule of studies on sentiment analysis performed on ChatGPT-themed tweets extracted after its launch (Haque et al., 2022; Korkmaz et al., 2023; Li et al., 2023).

Social Influence Theory (SIT) (Kelman, 1958) presents that social influence occurs when people alter their opinions, attitudes and behaviour (actions) in response to the influence of others. It was decided to use SIT for this study as it would allow us to understand the reasons for adopting and acceptance of technologies. It would make it possible to predict and comprehend how a new technology, like ChatGPT, would gain popularity using social media content. SIT identifies two types of social influence: normative influence and informational influence (Kelman, 1958; Snijders & Helms, 2014). Individuals adapt to a group's norms or expectations to achieve acceptance or approval, whereas informational influence occurs when individuals embrace a group's views or behaviour because they feel the group has better knowledge or expertise.

Different individuals respond to influences from others due to various reasons. There are three levels of influence (or reasons to be influenced); namely, compliance, identification and internalization (Kelman, 1958; Wang et al., 2013). First, compliance (lowest level of influence) refers to when an individual changes to the induced behaviour to satisfy, get the acceptance of others or avoid punishments. For example, a person may participate in online political activities to gain the acceptance of friends (David et al., 2019). Second, identification refers to when an individual adopts the opinions and actions of a group or an individual associated with the influencer. For example, a person may reduce single-use plastic as the induvial or group he or she admires promotes sustainable practices. Finally, internalization (highest level of influence) refers to when individuals genuinely understand the messages passed by the influencers and integrate those things into their core identity. For example, a person may follow videos of a veganism expert on YouTube and over time, understand the benefits of those practices and incorporate them into his/her daily life (Davis et al., 2019). In the social media context, various individuals can play a crucial role in shaping the opinions and behaviours of others on the use of ChatGPT for various domains including innovations, regulations of digital technologies and governance. Thus, in this study, the SIT has been employed to analyze how individuals are influenced by other people's opinions, attitudes, and behaviours.

Relationship between Subjective Norm and Social Influence

As per SIT, the subjective norm can be referred to as "perceived social pressure to perform or not to perform a behaviour" (Varnali & Gorgulu, 2015). The influence can be applied by the family, peers, friends, or the society at large. It indicates that when a person's friends in a social network engage in particular behaviours or express particular viewpoints, that person will likely engage in those same behaviours or have similar viewpoints as well. For instance, it has been found that a user's political involvement on social media and political knowledge increases when their friends participate and share in political activity on social media (David et al., 2019; Vitak et al., 2011). The influence type compliance described in SIT (Kelman, 1958) can be considered to occur due to subjective norms.

When considering the adoption of a technology like ChatGPT, subjective norms would play a key role in spreading the details of application scenarios and their benefits. If a person has a higher number of followers, that means they would try to use and share the details of the technology with peers. On the other hand, if a person leaves the community then the person would be losing the followers and the influence on others. Therefore, it is possible to hypothesise as follows in the context of the use of ChatGPT,

H1: The subjective norm of a Twitter user positively affects the social influence on readership

Relationship between Social Identity and Social Influence

Social identity can be described as the way an individual categorizes himself/herself and others based on characteristics and group memberships (Davis et al., 2019; Varnali & Gorgulu, 2015). For instance, it might be determined according to a person's gender, country, profession, or even affiliation with a fan club. Here, a person may try to improve their self-image based on their personal identity (maybe by affiliating with successful groups) (Zhou, 2011). This can be identified as part of internalization in SIT, where people accept information passed to others if they consider that information to be accurate and insightful (Varnali & Gorgulu, 2015). Thus, it can lead to higher social influence.

If a person's identity has been verified, it indicates the perceived support for the behaviour. Also, a prior study indicates that identity verification generates positive emotions among the viewers (Davis et al., 2019). Furthermore, verification of the identity of a person increases the acceptance by others. Prior studies have indicated that trust plays a key role in users' attitudes towards communication, content sharing and establishing new interactions (Abbas Naqvi et al., 2020; Abu-Salih et al., 2020; Oldeweme et al., 2021). Especially when considering the use of ChatGPT people will tend to follow content that is from verified accounts as they are of high quality or correct. Moreover, it would be good to investigate how trustworthiness impacts user influence (Abu-Salih et al., 2020). Therefore, it is possible to hypothesise as follows in the context of the use of ChatGPT,

H2: The strong social identity of a user positively affects their social influence on readership

Relationship between Social Media Commitment and Social Influence

Individuals who have been committed to social media would have greater social capital. There would be strong ties among the individuals who show similarity based on demographics, opinions and interests and thus would be able to create social cohesion and a community bond (Abu-Salih et al., 2020). Further, being on social media platforms for a long time facilitates meeting individuals with diverse backgrounds and opinions (Venkatesan et al., 2021).

Prior studies have identified that veteran users contribute more to online platforms like Twitter compared to new users (Cho & Wash, 2021). They would have more expertise on what has been discussed previously. The new users rather play a passive role to familiarise themselves with the platform and gather information and feedback from others (Lampe & Johnston, 2005). In online communities there are around 80-90% of "lurkers" and they would only consume information without contributing to the content (Tedjamulia et al., 2005). Furthermore, there is a possibility of new members being ignored and their posts not being commented on (Lampe & Johnston, 2005). Thus, it is possible to state that experienced users would be in a better position to apply higher influence on Twitter. That is, their tweets would be much read by the others. Therefore, it is possible to hypothesise as follows in the context of the use of ChatGPT,

H3: The strong social media commitment of a user positively affects their influence on readership

Relationship between Sentiment of the Content and Social Influence

Previous research has indicated that the substance of a tweet has a significant impact on social influence, regardless of a user's number of followers (Barbagallo et al., 2012; Pivecka et al., 2022). They have identified that the sentiment of a post is a key aspect of the content. Furthermore, several studies have identified that positive posts are self-promoting and can be influential (Alizadeh et al., 2020; Berger & Milkman, 2010). On the other hand, some studies have shown that negative sentiments have a high impact on social influence (Barbagallo et al., 2012). Therefore, it is possible to hypothesise as follows in the context of the use of ChatGPT,

H4: The positive sentiment of a tweet user positively affects their influence on readership

Domain-based Activity

It has been indicated that veteran participants are more active and popularize a particular topic (Tedjamulia et al., 2005). As per the initial analysis, several most commonly discussed domains have been identified (explained in a later section). As per the topic modelling approach domains like healthcare, education, cybersecurity, governance and robots are selected for further analysis. Thus, there may be certain areas that would be more interesting than others and can be used to influence the readers (Alizadeh et al., 2020). Furthermore, it has been identified that not all the content by a user has a similar influence. The social influence varies based on a topic (Hamzehei et al., 2017). For example, topics have been identified based on the pre-labelled hashtags (Montangero & Furini, 2015; Xiao et al., 2014). Similarly, there is a study being done to develop a topic-based social influence measure (Hamzehei et al., 2017). Therefore, it is possible to hypothesise as follows in the context of the use of ChatGPT,

H5: The use of domain-based discussions by the user positively affects their influence on readership

The research model of this study is presented in Figure 1.

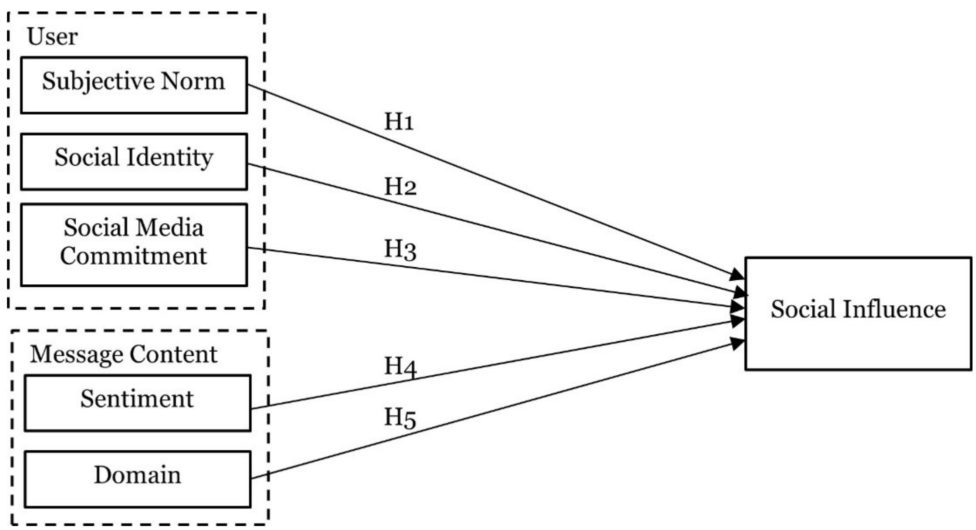


Figure 1. Research Model

Operationalization of the Measures

Recently Twitter introduced the view counts to indicate the amount of times a tweet has been viewed. Previously, to measure social influence studies have used the number of retweets, mentions, likes and comments for a post (Gong et al., 2021). Viewer means "an individual who takes a passive interest in the conversation" (Tinati et al., 2012). In Edelman's Topology of Influence (TOI) viewer is also considered as they gather, learn and share a lot of information in their offline network even though they do not actively participate in the discussion (Tinati et al., 2012). Since a person who generates can be considered an influencer if there are a lot of views, it was decided to use the view counts. Thus, the dependent variable social influence is measured using view count. When observing the posts related to ChatGPT, it is noted that the users have not shared multiple contents and they have not retweeted many of the tweets. Thus, as an indicator of the social influence view count is used. Many users use Twitter mainly as a way to get access to news and the latest updates. They mainly read the existing content rather than sharing new content about their daily life. Especially since the users have shared their personal concerns and reactions, limited retweets are observed. In this study, the View count is used as the dependent variable.

Follower count is used to represent the independent variable subjective norm. Followers refer to accounts that have subscribed to follow the user's content and be updated. That is, users can follow tweets from interested users. The content that is posted by the followers would be displayed on the feeds of the followers.

The number of followers is an indicator of the popularity of the users(Lerman & Ghosh, 2010). If a user has a high number of followers that means they are more influential (Bakshy et al., 2011). They have identified that word-of-mouth diffusion can only be achieved by having a large number of followers. In a similar line, it can be stated that a higher number of followers would have a positive relationship in the spread of ChatGPT user experiences related to the topics discussed.

Whether the user has a verified account is used to represent the independent variable social identity of the user account. A verified account means that the account has been confirmed as authentic by the providers of the platform. The account verification makes certain that the content is authentic and credible, as it will be shared by public figures and organizations (Tinati et al., 2012). Thus, it is possible to state that having a verified account (strong identity) would have a greater social influence as users consider the information to be truthful (Varnali & Gorgulu, 2015).

The independent variable social media commitment of the user is measured using the age of the Twitter account. How long the account has been existing is an indication of the level of experience of the user on social media platforms. For instance, users who have been on Twitter for a long period indicate that the user is committed to using Twitter (Sundar, 2008).

Furthermore, the content of the tweet is analysed based on the sentiment score (positive, negative, and neutral) which is mentioned as sentiment in the research model and the domain activity is measured using the topics discussed. As mentioned above, it was decided to consider healthcare, education, cybersecurity, governance and robotics.

As control variables, the number of media content shared and status count would be considered (Venkatesan et al., 2021). If Twitter shares more media content then there is a high possibility that many users will view the post (Abu-Salih et al., 2020). Furthermore, a user may be regularly updating their statuses which would be an indication of their frequent interaction with the social media platform.

The lag effect is not considered as almost all the users have posted only once about chatGPT. Since users have not discussed the topic earlier it is not possible to consider the impact of prior posts on the current post.

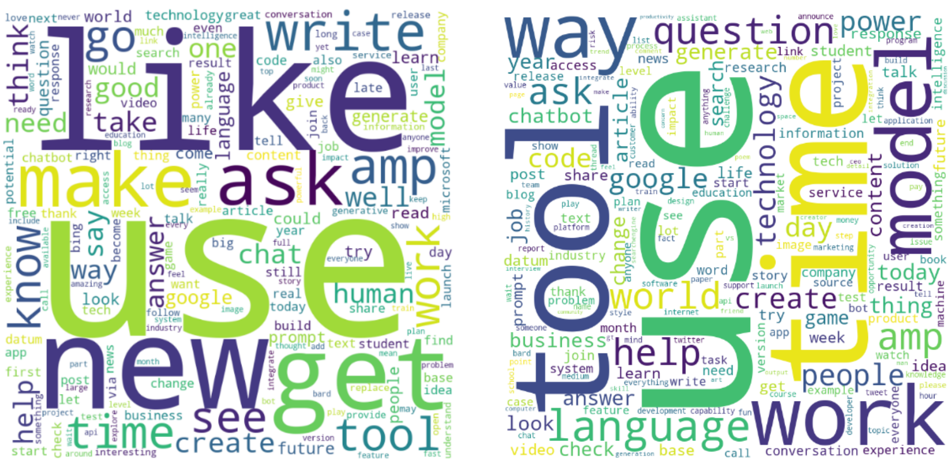
Methodology

Data Description

The study was carried out using tweets related from 1st January 2023 to 31st March 2023. The tweets were extracted based on the keywords "#chatgpt", "#gpt3", and "#gpt4" in English using the Snscrape (JustAnotherArchivist, 2018) Python library was used. The text extraction and processing were done using Jupyter Notebook and Google Colab Pro with Python 3. Originally there were 402,965 tweets with 31 attributes for the chosen period. The dataset underwent text normalization in which all links, mentions (e.g. @users), hashtags, unknown signs, accented characters, and emojis were removed. Moreover, all the text was converted to lowercase and expanded contractions as it would allow the models used in the study to focus on the content of the text rather than on the casing. It would make it easier to compare and analyse the text. Furthermore, duplicate tweets and tweets with less than 3 words were removed from the dataset. Subsequently, there were 332, 980 tweets for further analysis. The clean dataset 1 was prepared based on the pre-processing step mentioned above since it is important to maintain the original structure of text for transformer models (Egger & Yu, 2022).

The clean dataset 2 was prepared by further processing the clean dataset 1. Standard stop words were removed using NLTK, while words like "no" and "not" were whitelisted (without removing them) from the dataset as they would be required for the sentiment analysis. Furthermore, several common words such as "chatgpt", "gpt", "openai", and "ai" were removed as these terms were used in all the tweets to refer to ChatGPT. Besides, commonly occurring bi-grams were evaluated as "artificial intelligence" and "chat gpt" and they were removed from the text as they are highly frequent terms in all the content and will not generate any value in the analysis. The word cloud (Figure 2) generated for clean dataset 2 using a count vectorizer with unigram is shown below. Here, the words in less than 1% of tweets (rare words) and more than 75% of the tweets (most common words) were not considered in creating the word cloud in Figure 2. It was decided to remove rare and common words as (1) word cloud can be overcrowded and would be less visually appealing; (2) rare words might not contribute significantly to overall analysis and (3) common words might contain pronouns and other frequently occurring words that cannot be used to differentiate the tweets.

As can be seen in Figure 2, the most common terms such as make, like, and write are used indicating that people are mainly talking about their user experiences. Even though the data collection started a month after the service's inception, this shows that Twitter users continue to discuss its potential uses and preferences. The appearance of the word "use" signifies that many of the tweets are related to the sharing of experience on using GPT-3 and ChatGPT. The word "like" indicates that the users like the results that they received from ChatGPT. Then, the word "get" indicates the user experience of getting the GPT-3 or ChatGPT using subscription plans. The word "new" specifies the information shared related to the latest versions and updates of ChatGPT. The word "ask" mentions the user experience related to responses received for the questions asked. A similar, kind of word cloud (done without normalization of data) is obtained in prior literature too (Korkmaz et al., 2023). It was further, explored considering only the nouns, to identify specific areas the Twitter users have discussed. For that, words with 'NN' using the Part of Speech (POS) tagger in NLTK were used. Similar, results were obtained in both approaches where the user shared their personal experiences, use cases and comparisons with other similar tools.



(a) (b)

Figure 2. Wordcloud (a) Using Clean Dataset 2 and (b) Using Only Nouns

Topic Modelling

Topic modelling was applied to identify key topics being discussed in the large dataset related to ChatGPT. It was decided to use topic modelling as (a) topic modelling is the most frequently used technique to extract features from a large text dataset, (b) it is useful to identify hidden topical patterns and (c) useful for clustering tweets based on the topics (Shi et al., 2015). BERTTopic (Grootendorst, 2022), which is a topic modelling technique built upon BERT (Bidirectional Encoder Representations from Transformers) embeddings. (Egger & Yu, 2022; Sánchez-Franco & Rey-Moreno, 2022). It was decided to BERTopic over LDA as BERTopic (a) deals with complex document structures, (b) captures contextual semantics (compared to dependency on the bag of words by LDA), (c) does not require providing the number of topics, and (d) provides human-interpretable topic labels (Egger & Yu, 2022; Grootendorst, 2022). Clean dataset 1 was used as it is important to maintain the sentence structure for BERT-based models (Sánchez-Franco & Rey-Moreno, 2022). Uniform Manifold Approximation and Projection (UMAP) and Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) model with count vectorizer model have been used for feature engineering and dimensional reduction. BERTopic uses traditional features like word frequency counts (count vectorizer) along with the BERT embeddings (compact embedding model of BERT known as Paraphrase-MiniLM-L12-v2). In this case, UMAP was utilized to reduce the high-dimensional BERT embeddings to a lower dimension, making the data more manageable for clustering (Ghojogh et al., 2023). The 'curse of dimensionality problem' can be effectively handled using the UMAP model while keeping the local and global structures present in the datasets. Afterwards, a clustering algorithm known as HDBSCAN (Campello et al., 2015) was applied. HDBSCAN can automatically detect the number of clusters without doing manual parameter tuning, preserve the global structure of the dataset, and handle outliers as noise. To assign topic names to clusters of tweets, the class-based Term Frequency-Inverse Document Frequency (c-TF-IDF) was used (Grootendorst, 2022).

In Table 2, the categories used for the analysis are presented. The tweets in clean dataset 2 have been split into 87 topics. From the whole dataset 151, 243 tweets have not been assigned a topic. As per the hierarchical clustering performed certain topics are linked. Thus, they can be grouped. The group is mentioned in Table 2 as a Domain.

Domain	Topic No	Keywords
Productivity	1	Productivity, innovation, worker, developer, hire
	2	Chatbots, conversational, personality, plugin, slack
Education	3	student, education, teacher, educator, classroom
	56	Ban, cheat, school, cheating, student
Healthcare	5	brain, creativity, healthcare, doctor, patient
Cybersecurity	8	security, cybersecurity, malware, privacy, hacker
	11	Ethical, misinformation, truth, hallucination, moral
Governance	12	Bias, political, racist, politician, liberal
Robotics	24	Robot, robotic, robots, humanoid, overload

Table 2. Representative Topics within Each Domain

Sentiment Analysis

For the sentiment analysis, clean dataset 2 was used. The VADER which is a lexicon and rule-based sentiment analysis tool was applied to the dataset. Out of the 332980 tweets, there were 189614 positive sentiment tweets, 90579 neutral sentiment tweets, and 52787 negative sentiment tweets. As illustrated in Figure 3, there are many positive tweets compared to negative tweets.

It is obvious considering the advanced capabilities and use cases of ChatGPT, there would be highly positive opinions among the users. Also, it can be observed that there is a peak in the trend line on 15th March as on 14th March GPT-4 was released and many users have shared their opinion.

Analysis of Data

The data related to topics of productivity, healthcare, education, cybersecurity, governance and robotics were considered. The topics for further analysis were decided based on the research publications related to ChatGPT and the scope of the conference track. There were 43, 367 tweets belonging to these topics and they were used for further analysis. The other topics were not considered as they are not part of the scope of this paper. In the section below, descriptive statistics with correlations are given.

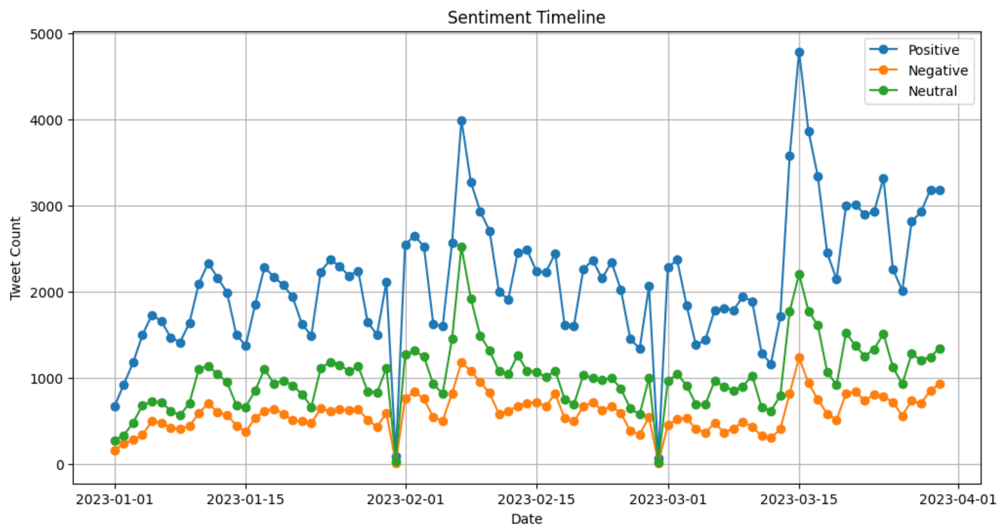


Figure 3. Twitter Sentiment Timeline

Descriptive Statistics and Correlations

Table 3 shows the descriptive statistics and correlation results. The majority of the correlations between variables are statistically significant at p <0.001. Multicollinearity may not be an issue because the correlations between the variables are less than 0.8 (threshold value) (Gujarati et al., 2012).

	Mean	SD	VC	D0	D1	D2	D3	D4	D5	V	Age	FC	S	SC
VC[a]	2.05	0.65	-											
D0	0.44	0.50	-0.08	-										
D1	0.19	0.39	-0.02	-0.43	-									
D2	0.19	0.39	0.16	-0.43	-0.24	-								
D3	0.09	0.28	-0.02	-0.28	-0.15	-0.15	-							
D4	0.05	0.23	-0.04	-0.21	-0.12	-0.12	-0.07	-						
D5	0.03	0.18	-0.01	-0.16	-0.09	-0.09	-0.06	-0.04	-					
V	0.05	0.22	0.29	-0.01	-0.01	0.00	0.03	-0.02	-0.01	-				
Age	0.82	0.39	0.13	-0.05	-0.02	0.11	0.01	-0.05	-0.02	0.16	-			
FC[a]	2.83	1.04	0.50	-0.06	-0.02	0.10	0.04	-0.05	-0.02	0.40	0.48	-		
S	0.26	0.42	0.03	0.08	0.08	0.04	-0.16	-0.15	-0.03	-0.02	0.01	-0.01	-	
SC[a]	3.54	0.94	0.18	-0.08	-0.01	0.05	0.05	0.04	0.00	0.25	0.49	0.74	-0.08	-
MC[a]	2.52	1.00	0.28	-0.06	-0.02	0.06	0.05	0.00	0.00	0.28	0.39	0.72	-0.05	0.82

Note. N= 43, 367; SD = Standard deviation; VC = View Count; D0 = Productivity; D1 = Healthcare3 D2 = Education; D3 = Cybersecuirty; D4 = Governance; D5 = Robotics; V = Is verified account; Age = Age of the account; FC = Followers count; S = Sentiment score; SC = Statuses score; MC = Media count; Correlations were significant at p<0.05 (2-tailed) except the underlined. [a]Log transformed variable.

Table 3. Descriptive Statistics and Correlation Matrix

Moreover, the Variance Inflation Factor (VIF) was performed (Table 3). Domain = Productivity has been removed from further analysis due to the high VIF value. VIF values range between 1.04 and 2.86 (VIF<3) for the dependent variables except for the control variables. That is there is no significant issue of multicollinearity (Amick et al., 1974; Fox, 1991).

Hypothesis Testing

Multiple linear regression was done using the SPSS statistical software tool to analyse whether the hypothesis was satisfied. As per the results indicated in Table 4, it is identified that the subjective norm (follower count) has a positive relationship with social influence (0.746, p<0.001). That means H1 is satisfied. Also, having a social identity (verified account = Yes) has a positive impact on social influence (0.078, p<0.001). Thus, H2 is also satisfied. There is a negative relationship between social media commitment (account age) and social influence (-0.073, n.s.). Thus, H3 is not satisfied. Further, the polarity (sentiment) score has a positive relationship with social influence (0.010, p<0.01). Thus, H4 is also satisfied. Moreover, the domains (topics) related to healthcare, education, governance and robotics have a significant positive influence on the user views. The domains healthcare, education, governance and robotics have beta coefficient values of 0.026 (p<0.001), 0.115 (p<0.001), 0.027 (p<0.001), and 0.015 (p<0.001) respectively. Thus, it is possible to state that H5(healthcare), H5(education), H5(governance) and H5(robotics) are satisfied. However, there is no positive impact of the cybersecurity topic on social influence (-0.007, n.s.). Thus, H5(cybersecurity) is not satisfied.

	Standardized Coefficients	95.0% Confidence Interval		VIF
	Beta	Lower Bound	Upper Bound	
(Constant)		1.652	1.694	
Statuses count[1a]	-0.443***	-0.318	-0.297	3.788
Media count[1a]	0.105***	0.060	0.077	3.323
Follower count [a]	0.746***	0.461	0.477	2.866
Verified account (=Yes)	0.078***	0.208	0.258	1.209
Age of the account [a]	-0.073***	-0.139	-0.108	1.389
Sentiment score	0.010**	0.003	0.028	1.066
Domain = Healthcare	0.026***	0.030	0.057	1.161
Domain = Education	0.115***	0.177	0.204	1.176
Domain = Cybersecurity	-0.007	-0.034	0.003	1.127
Domain = Governance	0.027***	0.056	0.102	1.105
Domain = Robotics	0.015***	0.028	0.084	1.042

Note. N= 43, 367; [1] control variable; * p < 0.05, ** p < 0.01, *** p < 0.001 (2-tailed). . [a]Log transformed variable

Table 4. Results of Hypothesis Testing

Discussion

The study was carried out using tweets related to ChatGPT collected from 1st January 2023 to 31st March 2023, using 402, 965 tweets. This paper provides a deeper understanding of the fundamentals by which social media may be effectively used to disseminate an idea in an influential manner. Initially, the text content of the tweets was analysed to categorise them into topics using deep learning. The topics identified were further analysed using a multiple linear regression model. According to the analysis, there is a positive relationship between a user's social identity and social influence for a tweet, as well as between subjective norm and social influence for a tweet and between content polarity and social influence for a tweet. Here, the number of views for a post is used as an indicator of the social influence of a tweet. Based on the findings, it is possible to state that there is a better social influence related to domains: education, healthcare, governance and robotics tweets shared about ChatGPT.

As per the number of research publications published and based on the tweets analysed, it is possible to state that many researchers and Twitter users are interested in those domains. Thus, tweets in such a domain can be very influential. Furthermore, considering the policy setting and government acceptance or banning of these technologies must have made Twitter users read such tweets. Similarly, Twitter users must have been interested in reading tweets related to innovations like robotics. Especially, if a certain user is interested in ChatGPT-like technology, then there is a high chance they are interested in robotics too.

The fact that there is no positive relationship as hypothesized between the age of the account and the social influence of a tweet deserves special attention. Even though it was hypothesised that more years of usage leads to active participation it may not be the case. There may be a possibility that older users are playing a passive idea related to a state-of-the-art technology like ChatGPT. New and younger users might have explored the platform and may be sharing their thoughts. Thus, there may be a negative relationship between the age of the account and the number of views as indicated in the hypothesis testing. This is similar in line with previous studies where young users are more likely to be enthusiastic towards new technological features (Sonderegger et al., 2016; Sundar, 2008). They are more likely to interact with the interactive features and share their opinions among their friends.

Further, the hypothesis of the relationship between the topic of cybersecurity and the social influence of a post is not satisfied. This may be because the topic of cybersecurity is about security and it may be a general topic related to any IT artefact and may not be generating the expected interest. Furthermore, among the early adopters of ChatGPT, as can be seen from the sentiment analysis, they have mainly discussed the positive points. As can be noted in Table 2, the main keywords for cybersecurity are related to the privacy of user data which would be a negative aspect related to the technology. Thus, having less negative sentiment can be a reason for not showing a significant relationship. However, with the use of technology, different countries and institutions may block it considering the privacy concerns. However, it needs to be further explored. Likewise, at the beginning productivity (domain) has been removed from further consideration due to multi-collinearity problems. This may be true, since for all the other domains (like healthcare and education), productivity is a common feature. Thus, it is not practical to study it as an independent variable.

Implications

It is possible to identify several theoretical implications. First, the use of view count to measure the social influence can be considered. Being a new feature introduced by Twitter it has not been used as a measure. However, to capture the role of passive readers who disseminate the ideas in offline networks it is important to consider the views. The view count would capture both active (before sharing would read the tweet) and passive readers rather than only considering the retweet count. Thus, in future, based on the findings of this study, the view count also can be used as a measure of social influence on social media sites.

Second, there is a dearth of studies analysing the influence of tweets on the dissemination of information related to a new technology. Thus, the use of the Social Influence theory to identify the factors relevant to making a social influence by a tweet can be instrumental in future research. Based on the findings it indicates that compliance and internalization play a significant role when increasing the readership for a new destructive technology like ChatGPT.

There are several practical implications of the study. First, the study analyses the discussion topics related to ChatGPT happening on Twitter. As per the word cloud drawn, it can be identified that the key terms are related to the use and applications of ChatGPT. Further, the analysis indicated that people mainly discuss the application of ChatGPT related to human-like text generation, responses to questions, and summarization. Also, healthcare and education are key discussion topics within the initial months after the launch of the platform. These findings indicate that the ChatGPT can be successfully used in these domains considering the positive sentiments available.

Second, beyond merely analysing the trends, this study analysed the factors enabling social influence (higher readership). Thus, it has been possible to identify that certain topics are more influential compared to others (Hamzehei et al., 2017). Through this study, it has been identified that certain discussion topics can attract many readers and thereby create a better social influence (Taecharungroj, 2023). For example, the education domain was identified to be a key discussion topic among Twitter users. Thus, it would be useful to go through these posts and develop relevant frameworks by the policymakers (Li et al., 2023).

Limitations and Future Work

When going through the results of this study, several limitations were identified. We have used data from a cross-section for a given period, however, there have not been previous posts by the user related to ChatGPT. To analyse the social influence it would have been better to consider the lag effect and by expanding the study period from 3 months to 12 months, there may be multiple posts by the same user. Furthermore, the analysis is limited to several domains that relate to society (healthcare, education, security, governance and robotics) (Dwivedi et al., 2023).

Thus, in future, it would be interesting to study the different discussion domains and analyse the level of impact on society (Haque et al., 2022) and users (Hamzehei et al., 2017). Furthermore, it would be useful to do an analysis of the influence of social media users and rank their level of influence (Bakshy et al., 2011; Venkatesan et al., 2021). Thus, various parameters related to the network can be further explored (Alizadeh et al., 2020).

Conclusion

This study is performed using tweets extracted related to ChatGPT after one month of its launch to understand the views of the public and the parameters that can create social influence. Based on the findings it is observed that users have mainly discussed their user experiences and how they like the new technology. They have mainly discussed the fields of education and healthcare. Additionally, it was discovered that subjective norms, identities, sentiments, and domains are crucial in promoting fresh debates and uses of cutting-edge technologies. These findings would thereby increase our understanding of social influence and user communication research.

References

Abbas Naqvi, M. H., Jiang, Y., Miao, M., & Naqvi, M. H. (2020). The effect of social influence, trust, and entertainment value on social media use: Evidence from Pakistan. *Cogent Business and Management*, *7*(1), 1723825. https://doi.org/10.1080/23311975.2020.1723825

Abu-Salih, B., Chan, K. Y., Al-Kadi, O., Al-Tawil, M., Wongthongtham, P., Issa, T., Saadeh, H., Al-Hassan, M., Bremie, B., & Albahlal, A. (2020). Time-aware domain-based social influence prediction. *Journal of Big Data*, *7*(1), 1–37. https://doi.org/10.1186/s40537-020-0283-3

Ahmad, A., Waseem, M., Liang, P., Fahmideh, M., Aktar, M. S., & Mikkonen, T. (2023). Towards Human-Bot Collaborative Software Architecting with ChatGPT. *ACM International Conference Proceeding Series*, *23*, 279–285. https://doi.org/10.1145/3593434.3593468

Alizadeh, M., Shapiro, J. N., Buntain, C., & Tucker, J. A. (2020). Content-based features predict social media influence operations. *Science Advances*, *6*(30). https://doi.org/10.1126/sciadv.abb5824

Androutsopoulou, A., Karacapilidis, N., Loukis, E., & Charalabidis, Y. (2019). Transforming the communication between citizens and government through AI-guided chatbots. *Government Information Quarterly*, *36*(2), 358–367. https://doi.org/10.1016/j.giq.2018.10.001

Bakshy, E., Mason, W. A., Hofman, J. M., & Watts, D. J. (2011). Everyone's an influencer: Quantifying influence on twitter. *Proceedings of the 4th ACM International Conference on Web Search and Data Mining, WSDM 2011*, 65–74. https://doi.org/10.1145/1935826.1935845

Barbagallo, D., Bruni, L., Francalanci, C., & Giacomazzi, P. (2012). An Empirical Study on the Relationship between Twitter Sentiment and Influence in the Tourism Domain. In *Information and Communication Technologies in Tourism 2012* (pp. 506–516). https://doi.org/10.1007/978-3-7091-1142-0_44

Berger, J., & Milkman, K. (2010). Social Transmission, Emotion, and the Virality of Online Content. *Wharton Research Paper*, 1–53. http://robingandhi.com/wp-content/uploads/2011/11/Social-Transmission-Emotion-and-the-Virality-of-Online-Content-Wharton.pdf

Biswas, S. S. (2023a). Role of Chat GPT in Public Health. In *Annals of Biomedical Engineering*. https://doi.org/10.1007/s10439-023-03172-7

Biswas, S. S. (2023b). Potential Use of Chat GPT in Global Warming. In *Annals of Biomedical Engineering* (pp. 1–2). Springer. https://doi.org/10.1007/s10439-023-03171-8

Campello, R. J. G. B., Moulavi, D., Zimek, A., & Sander, J. (2015). Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Transactions on Knowledge Discovery from Data*, *10*(1). https://doi.org/10.1145/2733381

Cho, J., & Wash, R. (2021). How Potential New Members Approach an Online Community. *Computer Supported Cooperative Work: CSCW: An International Journal*, *30*(1), 35–77. https://doi.org/10.1007/s10606-020-09390-0

David, C. C., San Pascual, R. S., & Torres, E. S. (2019). Reliance on Facebook for news and its influence on political engagement. *PLoS ONE*, *14*(3). https://doi.org/10.1371/journal.pone.0212263

Davis, J. L., Love, T. P., & Fares, P. (2019). Collective Social Identity: Synthesizing Identity Theory and Social Identity Theory Using Digital Data. *Social Psychology Quarterly*, *82*(3), 254–273. https://doi.org/10.1177/0190272519851025

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (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. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Egger, R., & Yu, J. (2022). A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts. *Frontiers in Sociology*, *7*, 886498. https://doi.org/10.3389/fsoc.2022.886498

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*. https://doi.org/10.1080/14703297.2023.2195846

Ghojogh, B., Crowley, M., Karray, F., & Ghodsi, A. (2023). Uniform Manifold Approximation and Projection (UMAP). In *Elements of Dimensionality Reduction and Manifold Learning* (pp. 479–497). Springer International Publishing. https://doi.org/10.1007/978-3-031-10602-6_17

Gong, Q., Chen, Y., He, X., Xiao, Y., Hui, P., Wang, X., & Fu, X. (2021). Cross-site prediction on social influence for cold-start users in online social networks. *ACM Transactions on the Web*, *15*(2), 22. https://doi.org/10.1145/3409108

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. http://arxiv.org/abs/2203.05794

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy. *IEEE Access*, *11*, 80218–80245. https://doi.org/10.1109/ACCESS.2023.3300381

Hamzehei, A., Jiang, S., Koutra, D., Wong, R., & Chen, F. (2017). Topic-based social influence measurement for social networks. *Australasian Journal of Information Systems*, *21*. https://doi.org/10.3127/ajis.v21i0.1552

Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., & Ahmad, H. (2022). *"I think this is the most disruptive technology": Exploring Sentiments of ChatGPT Early Adopters using Twitter Data*. https://openai.com/

Howard, A., Hope, W., & Gerada, A. (2023). ChatGPT and antimicrobial advice: the end of the consulting infection doctor? In *The Lancet Infectious Diseases* (Vol. 23, Issue 4, pp. 405–406). Elsevier Ltd. https://doi.org/10.1016/S1473-3099(23)00113-5

Hu, K. (2023). ChatGPT sets record for fastest-growing user base - analyst note. *Reuters*, 1–10. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

JustAnotherArchivist. (2018). *snscrape · PyPI*. https://pypi.org/project/snscrape/

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. In *Learning and Individual Differences* (Vol. 103, p. 102274). JAI. https://doi.org/10.1016/j.lindif.2023.102274

Kelly, S. M. (2023). *CHATGPT passes exams from law and Business Schools*. CNN Business. https://edition.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html

Kelman, H. C. (1958). Compliance, identification, and internalization three processes of attitude change.

Journal of Conflict Resolution, *2*(1), 51–60. https://doi.org/10.1177/002200275800200106

King, M. R. (2023). A Conversation on Artificial Intelligence, Chatbots, and Plagiarism in Higher Education. In *Cellular and Molecular Bioengineering* (Vol. 16, Issue 1, pp. 1–2). Springer. https://doi.org/10.1007/s12195-022-00754-8

Koetsier, J. (2023). *GPT-4 Beats 90% Of Lawyers Trying To Pass The Bar*. Forbes. https://www.forbes.com/sites/johnkoetsier/2023/03/14/gpt-4-beats-90-of-lawyers-trying-to-pass-the-bar/?sh=308a3c243027

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. In *RELC Journal*. SAGE Publications Ltd. https://doi.org/10.1177/00336882231162868

Korkmaz, A., Akturk, C., & Talan, T. (2023). Analyzing the User's Sentiments of ChatGPT Using Twitter Data. *Iraqi Journal for Computer Science and Mathematics*, 202–214. https://doi.org/10.52866/ijcsm.2023.02.02.018

Lampe, C., & Johnston, E. (2005). Follow the (slash) dot: Effects of feedback on new members in an online community. *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work*, 11–20. https://doi.org/10.1145/1099203.1099206

Lerman, K., & Ghosh, R. (2010). Information contagion: An empirical study of the spread of news on Digg and Twitter social networks. *ICWSM 2010 - Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, 90–97. https://doi.org/10.1609/icwsm.v4i1.14021

Li, L., Ma, Z., Fan, L., Lee, S., Yu, H., & Hemphill, L. (2023). *ChatGPT in education: A discourse analysis of worries and concerns on social media*. http://arxiv.org/abs/2305.02201

Liebrenz, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating scholarly content with ChatGPT: ethical challenges for medical publishing. In *The Lancet Digital Health* (Vol. 5, Issue 3, pp. e105–e106). Elsevier Ltd. https://doi.org/10.1016/S2589-7500(23)00019-5

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries? In *Library Hi Tech News* (Vol. 40, Issue 3, pp. 26–29). Emerald Publishing. https://doi.org/10.1108/LHTN-01-2023-0009

Mijwil, M. M., Aljanabi, M., & ChatGPT. (2023). Towards Artificial Intelligence-Based Cybersecurity: The Practices and ChatGPT Generated Ways to Combat Cybercrime. *Iraqi Journal for Computer Science and Mathematics*, *4*(1), 65–70. https://doi.org/10.52866/ijcsm.2023.01.01.0019

Montangero, M., & Furini, M. (2015). TRank: Ranking Twitter users according to specific topics. *2015 12th Annual IEEE Consumer Communications and Networking Conference, CCNC 2015*, 767–772. https://doi.org/10.1109/CCNC.2015.7158074

Oldeweme, A., Märtins, J., Westmattelmann, D., & Schewe, G. (2021). The role of transparency, trust, and social influence on uncertainty reduction in times of pandemics: Empirical study on the adoption of COVID-19 tracing apps. *Journal of Medical Internet Research*, *23*(2), e25893. https://doi.org/10.2196/25893

Oviedo-Trespalacios, O., Peden, A. E., Cole-Hunter, T., Costantini, A., Haghani, M., Rod, J. E., Kelly, S., Torkamaan, H., Tariq, A., Newton, J. D. A., Gallagher, T., Steinert, S., Filtness, A., & Reniers, G. (2023). The Risks of Using Chatgpt to Obtain Common Safety-Related Information and Advice. *SSRN Electronic Journal*, *167*, 106244. https://doi.org/10.2139/ssrn.4370050

Patel, S. B., & Lam, K. (2023). ChatGPT: the future of discharge summaries? In *The Lancet Digital Health* (Vol. 5, Issue 3, pp. e107–e108). Elsevier Ltd. https://doi.org/10.1016/S2589-7500(23)00021-3

Pivecka, N., Ratzinger, R. A., & Florack, A. (2022). Emotions and virality: Social transmission of political messages on Twitter. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.931921

Qi, X., Zhu, Z., & Wu, B. (2023). The promise and peril of ChatGPT in geriatric nursing education: What We know and do not know. *Aging and Health Research*, *3*(2), 100136. https://doi.org/10.1016/j.ahr.2023.100136

Sallam, M. (2023). ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. In *Healthcare (Switzerland)* (Vol. 11, Issue 6, p. 887). Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/healthcare11060887

Sánchez-Franco, M. J., & Rey-Moreno, M. (2022). Do travelers' reviews depend on the destination? An analysis in coastal and urban peer-to-peer lodgings. *Psychology and Marketing*, *39*(2), 441–459. https://doi.org/10.1002/mar.21608

Shi, Z. M., Lee, G. M., & Whinston, A. B. (2015). Toward a better measure of business proximity: Topic modeling for industry intelligence. *MIS Quarterly*, *40*(4), 1035–1056. https://doi.org/10.25300/misq/2016/40.4.11

Snijders, R., & Helms, R. (2014). Analyzing social influence through social media, a structured literature review. *Proceedings of the 7th IADIS International Conference Information Systems 2014, IS 2014*, 3–10. https://research.ou.nl/en/publications/analyzing-social-influence-through-social-media-a-structured-lite

Sonderegger, A., Schmutz, S., & Sauer, J. (2016). The influence of age in usability testing. *Applied Ergonomics*, *52*, 291–300. https://doi.org/10.1016/j.apergo.2015.06.012

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. *Digital Media, Youth, and Credibility*, 73–100. https://doi.org/10.1162/dmal.9780262562324.073

Surameery, N. M. S., & Shakor, M. Y. (2023). Use Chat GPT to Solve Programming Bugs. *International Journal of Information Technology and Computer Engineering*, *3*(1), 17–22. https://doi.org/10.55529/ijitc.31.17.22

Taecharungroj, V. (2023). "What Can ChatGPT Do?" Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data and Cognitive Computing*, *7*(1), 35. https://doi.org/10.3390/bdcc7010035

Tedjamulia, S. J. J., Olsen, D. R., Dean, D. L., & Albrecht, C. C. (2005). Motivating content contributions to online communities: Toward a more comprehensive theory. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 193. https://doi.org/10.1109/hicss.2005.444

Tinati, R., Carr, L., Hall, W., & Bentwood, J. (2012). Identifying communicator roles in Twitter. *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web Companion*, 1161–1168. https://doi.org/10.1145/2187980.2188256

Tinholt, D., et al. (2017). Unleashing the potential of Artificial Intelligence in the Public Sector. In *Capgemini Consulting*. https://www.capgemini.com/consulting/wp-content/uploads/sites/30/2017/10/ai-in-public-sector.pdf

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 1–24. https://doi.org/10.1186/s40561-023-00237-x

Varnali, K., & Gorgulu, V. (2015). A social influence perspective on expressive political participation in Twitter: the case of #OccupyGezi. *Information Communication and Society*, *18*(1), 1–16. https://doi.org/10.1080/1369118X.2014.923480

Venkatesan, S., Valecha, R., Yaraghi, N., Oh, O.-O., & Rao, H. R. (2021). Influence in Social Media: An Investigation of Tweets Spanning the 2011 Egyptian Revolution. *MIS Quarterly*, *45*(4), 1679–1714. https://doi.org/10.25300/misq/2021/15297

Vitak, J., Zube, P., Smock, A., Carr, C. T., Ellison, N., & Lampe, C. (2011). It's complicated: Facebook users' political participation in the 2008 election. *Cyberpsychology, Behavior and Social Networking*, *14*(3), 107–114. https://doi.org/10.1089/cyber.2009.0226

Wang, Y., Meister, D. B., & Gray, P. H. (2013). Social influence and knowledge management systems use: Evidence from panel data. *MIS Quarterly: Management Information Systems*, *37*(1), 299–313. https://doi.org/10.25300/MISQ/2013/37.1.13

Xiao, F., Noro, T., & Tokuda, T. (2014). Finding news-topic oriented influential twitter users based on topic related hashtag community detection. *Journal of Web Engineering*, *13*(5–6), 405–429. https://doi.org/10.5555/2685129.2685132

Zhou, T. (2011). Understanding online community user participation: A social influence perspective. *Internet Research*, *21*(1), 67–81. https://doi.org/10.1108/10662241111104884