How do Twitter Users Perceive ChatGPT?

Framing of ChatGPT on Twitter

Luiza Julia Świerzawska

7 July 2023

Utrecht University, the Netherlands

Applied Data Science

Word count (excl. references and appendix): 7824

Supervisor: Dr. Dennis Nguyen

Abstract

The launch of ChatGPT, an AI chatbot, has ignited a profound discussion on the future of artificial intelligence. With its impressive capabilities and wide-ranging applications, ChatGPT has garnered both praise and criticism, attracting 1 million users within only 5 days. While scholars and experts delve into the long-term implications of these technologies, people have shared their experience with the newest technologies on social media, including Twitter, a platform known for its microblogging and widespread discussions between users. Recognizing the significance of this recent development in the field of large language models, this study sets out to explore the perceptions of ChatGPT by analysing a vast dataset of 2.6 million Twitter posts using natural language processing techniques. By examining emphasis frames, sentiments, and the key actors involved, this research aims to shed light on the nature of the debates surrounding ChatGPT.

Keywords: Artificial Intelligence, Technology, ChatGPT, Chatbots, Twitter, Framing

Table of Contents

Introduction

Prominent historian Yuval Noah Harari recently penned an article for The Economist, in which he made a claim**:** "AI has hacked the operating system of our civilization." (Harari, 2023). Harari's assertion pointed towards AI's capabilities to manipulate and generate language and aimed to portray the recent advancements made in the realm of large language models (LLMs). On the 30th of November 2022, OpenAI released their AI chatbot, ChatGPT, inducing a paradigm shift in the discourse surrounding artificial intelligence. Within only 5 days, the new AI tool attracted 1 million users, sparking intense debates across various fields and marking another step towards artificial general intelligence (AGI) (Taecharungroj, 2023).

Traditionally, chatbots utilised natural language processing (NLP) methods which allowed them to understand patterns in human language and map users' queries to predefined responses available in the systems (Taecharungroj, 2023; Dwivedi et al., 2023). ChatGPT (Generative Pre-Training Transformer) is a step further. It deploys a transformer algorithm, a type of neural network, which is able to analyse and comprehend large amounts of textual data and later generate output that resembles human conversation (Dwivedi et al., 2023; Chadha, 2023). ChatGPT's coherent and plausible-sounding responses in conjunction with its easy-to-use interface (Taecharungroj, 2023), has induced a lot of enthusiasm among the public, quickly becoming an essential tool in a wide variety of tasks, ranging from programming to writing emails and producing bedtime stories.

Despite the widespread enthusiasm surrounding ChatGPT, multiple scholars have raised concerns about limitations and broader implications of LLMs (Dwivedi et al., 2023; Taecharungroj, 2023; Talboy & Fuller, 2023). Given the widespread use of ChatGPT, these fears are becoming increasingly salient. Open AI has come under fire by authorities and organisations as many have questioned the chatbot's adherence to privacy regulations and the soundness of its output. Italy was the first Western country to ban the latest version of ChatGPT (McCallum, 2023). This decision has also been mirrored by individual schools (ibid.). Scholars and authorities have called for more attention to educating users on the mechanisms behind AI systems like ChatGPT and developing appropriate regulations that will mitigate its potential malfunctions (Dwivedi et al., 2023; McCallum, 2023). On the 22nd of March 2023, more than 1000 technology experts, including entrepreneurs and scholars, signed an open letter released by the Future of Life Institute, urging artificial intelligence labs to pause the developments of AI tools more advanced than GPT-4, the latest version of ChatGPT (Metz & Schmidt, 2023). The petition has attracted considerable attention from the media due to its alarming tone on the risks that these systems pose to society and the endorsement of prominent figures, like Elon Musk, who have added their signatures (ibid; Firth-Butterfield, 2023). Nevertheless, the letter has faced criticisms from many scholars, pointing out that instead of calling for concrete actions, the petition has only fueled the

‘hype’ around AI (Xiang, 2023). Emily M. Bender, a renowned scholar from the University of Washington, has emphasised that the open letter is missing the point (ibid.). She explains that the risks are not stemming from ‘too powerful AI’ but rather from the centralization of power among individuals, the perpetuation of oppressive systems, the negative impact on the information ecosystem, and environmental harms due to excessive utilisation of energy resources (Bender, 2023).

Given the considerable importance of the AI chatbot, this research attempts to illuminate the discussions on Twitter surrounding the deployment of ChatGPT by OpenAI. To attain this goal, it offers a critical-empirical analysis that incorporates appropriate NLP techniques to uncover key themes, sentiments and trends pertaining to this discourse as well as public figures engaged in these debates. With the use of multiple quantitative methods, this research contributes by providing an extensive analysis of a large corpus of data, including four months of online Twitter discussions. Studying Twitter, as a popular platform for sharing opinions and engaging in conversations with others, offers a good insight on how the public at large is perceiving ChatGPT. Despite the limitations of implemented NLP techniques, derived results offer valuable observations on the attitudes, concerns and opinions of the public towards current matters such as technological advancements.

Research Plan and Research Questions

The paper begins with establishing a comprehensive overview and evaluation of existing research related to ChatGPT, as well as other large language models. It aims to identify implications of these AI systems and fields where it has been especially contested. Later on, in the analysis section, the results will reveal the characteristics of Twitter debate surrounding OpenAI's chatbot and whether the online discussions align with the insights derived through the literature review. The analysis chapter is organised according to the three delineated subquestions:

1. What themes can be identified in the discourse on ChatGPT on Twitter?
2. What sentiments emerge from the tweets on ChatGPT? Does it change over time?
3. Who are the key actors in the Twitter discourse surrounding ChatGPT?

Having answered the three questions, this research aims to illuminate various aspects of the discourse surrounding ChatGPT on Twitter and its dynamics, eventually allowing us to answer the main research question: How do Twitter users perceive ChatGPT? The scope of this study is narrowed down to Twitter, a platform with around 500 million posts (tweets) being sent daily by people engaging in discussions and microblogging (Shepherd, 2023). Due to that, Twitter data has been widely used in a variety of studies for the purpose of investigating public perceptions (Taecharungroj, 2023).

Given the large size of the dataset that this research utilises, the implementation of computational methods is of great benefit. These techniques offer a way to efficiently analyse vast amounts of records from various sources, including social media (Breuer, 2020). This presents an opportunity for the academic community to unravel latent patterns and trends present in complex datasets as well as encourages the deployment of innovative tools to improve the breadth and depth of one's investigation. Following this logic, this thesis utilises an extensive dataset derived from Twitter to hopefully shed some light on the nature of the discourse surrounding ChatGPT and its framing. It enables a systematic and data-driven understanding of public opinions pertaining to the controversial AI chatbot, providing valuable insights into its overall impact in the online space.

Literature Review

The following section provides a critical evaluation of the academic literature on ChatGPT. So far, there are three main topics, or domains, that can be identified with respect to the existing literature. These are the chatbot's impact on education, the bias ingrained in its training dataset, and the consequences for the labour market. Additionally, this thesis utilises the concept of framing, which is briefly discussed at the beginning of this section.

Twitter Discourse and Framing

Data generated by Twitter users has become increasingly popular amongst researchers, opening up new avenues in analysing public opinions, sentiments and attitudes (Wicke & Bolognesi, 2020). There are countless examples of research, *inter alia*, tweets have been used in the context of measuring people's reactions towards crises, like COVID-19 pandemic, or understanding the platform's role in shaping public opinion. In the context of ChatGPT, analysing the discussions of Twitter users can unravel their approaches towards the AI chatbot and the focal points of these conversations. Taecharungroj (2023) emphasises that the launch of this AI system has caused a lot of hysteria and hype in the first weeks. He outlined multiple use cases where ChatGPT is used. Having a dataset that covers four months, this study can further investigate the applications and sentiments pertaining to this language tool.

Furthermore, the analysis of Twitter posts is often utilised in conjunction with the theory of framing (Burch et al., 2015). The practice of framing information entails a collection of certain elements from a perceived reality that are emphasised and together provide a specific interpretation or perspective on a topic (Stefanone et al., 2015; Burch et al., 2015; Entman, 1993). It is a cognitive process through which individuals shape their understanding of an issue. The utilisation of Twitter in the context of framing was described by Nisbet (2009) as a 'bottom up model of framing' which describes a situation in which individuals who generate online content contribute to the framing of issues (Burch et al., 2015). This thesis incorporates two types of framing, emphasis and valence framing. The former type entails selectively highlighting specific aspects of a situation. The presence of particular topics sets the scope of discussions, focusing people's attention on a subset of the aspects of an issue (Yang and Pan, 2016). The latter type refers to the emotionality of a message and, therefore, aims to discover people's sentiments towards issues (ibid; Stefanone et al., 2015). Discussing the results with respect to these two concepts sheds the light on how Twitter users construe the implications of ChatGPT.

Education

The role of ChatGPT has been especially contested in the academic setting and in a broader educational context. Transforming the way we obtain information and produce text, the AI chatbot is expected to revolutionise the traditional learning methods (Else, 2023). Some view it as an opportunity for schools

and universities to alter their priorities and incorporate AI into their domain (Dwivedi et al., 2023). Dwivedi et al. (2023) argue that the recent technological advancements allow these institutions to place more focus on skills that add value beyond AI such as critical thinking. Moreover, since ChatGPT is capable of assisting in writing various types of content, the ability to reach a broad audience with one's research findings is no longer determined by superior writing skills. The chatbot is expected to help surpass the language barriers and allow researchers to focus on, what according to Laumer (2023) is the main objective of academia, "(…) *asking thought-provoking questions and conducting research to find answers"* (ibid.). Due to that, students need to be educated on the functioning of increasingly omnipresent algorithms.

However, despite the potential benefits that ChatGPT brings into academia, multiple studies have confirmed that ChatGPT often fails to answer factual questions (Taecharungroj, 2023; Dwivedi et al., 2023). What scholars find even more problematic is the confident tone of its output, which creates an illusion of the answer being correct. This can have severe consequences, deceiving users and potentially amplifying errors such as misinformation (Talboy & Fuller, 2023). Given the cohesive and well-formed output that the chatbot provides, users are susceptible to accepting its answers straight away, further exacerbating misunderstandings surrounding global issues such as politics, medicine or science (ibid.). Additionally, Rana et al. (2023) highlight that a question worded differently might derive a very different answer. Therefore, scholars stress that addressing the reliability of ChatGPT's answers as well as educating users on the potential limitations of LLMs are of the utmost importance (Talboy & Fuller, 2023; Ray, 2023).

Bias

The most significant limitations of the AI chatbot stem from its training data. Importantly, ChatGPT's knowledge is derived from its pre-2021 database, and currently lacks the capability to browse the internet to fill potential knowledge gaps. Due to that, its output can be outdated. Even more problematic than the data scarcity is the bias ingrained in the training data. Trained on an enormous amount of text data, ranging from books to Wikipedia pages and news articles, ChatGPT is prone to acquiring any biases or inaccuracies present in these sources. Due to inherent human biases, it has a tendency to internalise certain stereotypes and subsequently express them in its responses which can be, unintentionally, offensive and harmful (Ray, 2023). Research conducted by Talboy & Fuller (2023) confirmed that, while still performing better than its rivals, ChatGPT is susceptible to embedded social biases relating to, for example, personal temperament and social interactions. At the same time, these harmful assumptions often relate to representation of different minority groups.

Moreover, by reinforcing pre-existing biases and beliefs, AI systems can exacerbate extremist views and contribute to the polarisation of society. Therefore, questions have been asked about the chatbot's

political leanings. Despite Open AI's efforts to make ChatGPT as neutral as possible and abstaining from political discussions, Rozado (2023) discerns that when asked normative questions about "what ought to be", ChatGPT expresses left-leaning political orientation. According to the author, it is an undesirable outcome, since AI systems should primarily adopt viewpoints that are substantiated by factual evidence and reasoning (Rozado, 2023). Overall, chatbots entering the political realm may pose several dangers. This includes the aforementioned spread of false information as well as manipulation of public opinion. According to a study featured in the MIT Technology Review, approximately 20% of Twitter posts related to the 2016 US elections were generated by bots and aimed to sway opinions (Yao, 2016). With current AI advancements, experts are asking how we can keep the political discourse "human-only" (Foggatt, 2023). The plausible tone of texts generated by more advanced chatbot, such as ChatGPT, could blur the line between their output and content written by people, further exacerbating the issue.

Labor Market and Automation

Furthermore, with the unstoppable technology developments, the labour market is one of the most critical areas. The fears of automation have a long history. Implementation of the newest technologies have caused disruptions by making some jobs no longer relevant and simultaneously necessitating the creation of new positions (Zarifhonarvar, 2023). However, the launch of ChatGPT has slightly changed the usual trajectory, putting even "safe" jobs at risk (Zarifhonarvar, 2023; Taecharungroj, 2023). AI coders and analysts, as well as many other white-collar workers, have been perceived as irreplaceable for a long time(Taecharungroj, 2023). With the next job revolution, leaders in both the public and business sectors are forced to rethink the future of the labour market in this rapidly changing landscape (ibid.). At the same time, it is equally important for the education sector to consider how to prepare the next generation for the fast-changing world, where different skills and knowledge will be needed (Zarifhonarvar, 2023). To emphasise the scale of this disruption, Taecharungroj (2023) finds that "the next evolution of jobs" is the main issue discussed in the Twitter early reactions to ChatGPT.

Methodology and Data

This section delineates steps that were taken to conduct the analysis of the Twitter posts relating to ChatGPT as well as explains the incorporated NLP methods contributing to this research.

Dataset

The dataset implemented in this research consists of 2635610 tweets in English extracted using the Twitter API, excluding retweets. After deletion of duplicates, it was reduced to 1979436 data points. Posts mentioning ChatGPT were selected spanning over 4 months, from the 1st of January 2023 till the 28th of April 2023. The dataset contains 32 columns, all of which are listed in Appendix 1. However, only a subset of these was used. The majority of the conducted analysis was performed on the 'body' column, containing the Twitter posts. The information included in the 'timestamp' was utilised to display any variation over time.

Methods

This study incorporates multiple NLP techniques to provide an extensive analysis of Twitter discourse on OpenAI's chatbot. Latent Dirichlet Allocation topic modelling (LDA TM) is an unsupervised machine learning method which represents documents as a collection of topics, allowing a researcher to uncover themes present in the text (Kulshrestha, 2019). It will therefore play a crucial role in unravelling the main emphasis frames in the Twitter discourse pertaining to ChatGPT. Each topic was manually labelled following the close reading of its corresponding tweets. Moreover, to enrich the analysis of the latent themes pertaining to the discourse on ChatGPT, Bidirectional Encoder Representations from Transformers (BERT) is applied, a different topic modelling method. Furthermore, to assess the sentiment of tweets, sentiment analysis using Valence Aware Dictionary and Sentiment Reasoner (VADER) was conducted. Every tweet was assigned a score between -1 (fully negative) to 1 (fully positive). In order to explore the key actors engaging in the debates on ChatGPT, Named Entity Recognition (NER) was employed.

Data Preprocessing

The tweets' content underwent multiple cleaning steps. This included tokenization, removal of punctuation, non-alphabetic, special characters and emojis, lemmatization as well as removing stopwords, which were extended based on the topic of this study. For LDA topic modelling, only nouns were included after extracting them with SpaCy.

Analysis

Most Frequent Hashtags

Identifying words that are common in the discourse surrounding ChatGPT can help to understand the domains in which the implications of this technology are mostly debated. It can facilitate our comprehension of the context in which users most frequently share their opinions and arguments. Several approaches were implemented to discern the main emphasis frames on Twitter discussions surrounding ChatGPT. The analysis begins with extracting the most frequent hashtags present in tweets on OpenAI's chatbot. Twitter's hashtags are a popular way of indicating keywords or a topic associated with the content of one's message (Hemphill et al., 2013). Therefore, displaying the list of the most used ones can already unravel some areas of discussions pertaining to ChatGPT.

Figure 1. Top 15 most frequent hashtags in the Twitter discourse on ChatGPT.

hashtag	count
#arbitrum	51356
#zenithswap	50509
#nft	25409
#crypto	21060
#crypto	15174
#airdrop	14153
#machinelearning	13053
#google	11957
#microsoft	10925
#chatbot	10914
#web3	9952
#generativeai	9697
#bitcoin	9343
#web3	8834
#news	8202

As listed above, the most frequent hashtags appearing in the analysed Twitter dataset relate to cryptocurrency and pinpoint an association between crypto-related matters and ChatGPT. These hashtags include #arbitrum, #zenithswap, #nft, #airdrop and #bitcoin, terminology belonging to the cryptocurrency field. There are numerous connections between ChatGPT and crypto trading, but the most popular one is the use of the chatbot as an assisting tool in trading strategies. The hashtag analysis also revealed the presence of companies like Google and Microsoft in the discussions pertaining to ChatGPT, stressing the importance of the AI chatbot within the tech industry as well as suggesting an important role for tech giants in the development of LLMs.

Topic Modeling – Latent Dirichlet Allocation (LDA)

Given that the framing of a certain issue can be determined through the selection of topics present in the discourse, this thesis utilises a popular method of topic modelling, LDA, to uncover latent themes

occurring in the discourse on ChatGPT. The technique generated 18 interpretable topics with a coherence score of 0.32. Each topic consisted of a set of representative keywords. Based on these keywords, each theme was given a label, which is an interpretation of the LDA results.

Table 1. Topics generated with LDA Topic Modeling from Tweets on ChatGPT.

Topic Label	**Keywords**	**Proportion of Tweets**
Applications in Work and Education	work, lot, education, version, hour, example, end, page, report, expert	13.02%
Chatbots and Product Development	datum, bot, product, stuff, step, computer, guy, challenge, info, creativity	3.53%
Applications in Content Creation	search, content, ai, writing, video, engine, business, website, paper, creator	6.06%
Web Development and Legal Considerations	internet, developer, access, api, course, law, exam, change, review, cost	2.67%
Opportunities in Generating Content	result, email, test, poem, reason, type, hype, moment, function, update	2.73%
Recent Developments in Tech Sector	day, tech, today, level, fact, ability, power, bit, service, data	4.65%
AI Tools	ai, tool, chatbot, technology, company, knowledge, project, research, software, intelligence	11.26%
Education and Learning Resources	article, app, school, teacher, image, mind, way, feature, minute, kid	3.65%
Implications of AI	essay, life, ai, money, artificialintelligence, art, machinelearning, piece, instruction, note	3.57%
Discussions and News on Technology	tweet, link, news, development, crypto, summary, team, resource, author, comment	3.75%

AI Applications and Language Processing	model, language, year, post, ai, blog, task, friend, process, voice	6.40%
AI for Digital Marketing in Gaming Industry	people, future, ai, game, marketing, world, talk, industry, impact, capability	6.89%
Skills and Productivity	time, human, idea, word, story, information, seo, person, skill, plan	6.29%
Utilisation of AI in Coding	code, student, learning, point, topic, book, list, machine, programming, coding	5.03%
Features and Use of AI Language Models	question, response, llm, thought, week, intelligence, experience, value, copy, query	5.09%
Utilisation of Chatbots	user, conversation, chat, source, month, issue, help, platform, form, hand	4.05%
Communication with Chatbots	thing, answer, prompt, use, problem, case, thank, question, solution, robot	7.21%
Text Generation for Professionals	text, thread, job, system, output, training, startup, line, input, advice	4.13%

Despite being challenging to interpret, the generated keywords can provide some insights on the Twitter discourse surrounding ChatGPT. What seems to be a prevalent pattern is that the AI chatbot has been widely discussed with respect to its broad range of applications. The most frequent theme identified by LDA in 13.02% of tweets, *Applications in Work and Education*, refers to a variety of discussions on how the newest LLMs' advancements will impact these two fields. Similar notions seem to be encompassed by the topic *Education and Learning Resources* which emphasises the presence of conversations regarding schooling. Twitter users are divided on whether the latest advancements in the LLM realm will benefit the educational sector or rather pose a threat. The close read tweets relating to education contained some sceptical opinions, viewing ChatGPT as a tool for cheating at school and spreading misinformation. Moreover, concerns were expressed about students of different ages not acquiring important skills due to frequent use of AI tools. On the other hand, some people perceive it as an opportunity to restructure the educational system in a way that it benefits from future technological advancements.

In a very similar manner, discussions on job markets questioned ChatGPT's impacts on people's careers. Users share their opinions on which jobs are at risk of being replaced and which will be in high demand. Similar trends were reflected in the academic literature on ChatGPT, where scholars are debating the unstoppable automation and its consequences. However, given the topics generated by LDA and close reading of the corresponding tweets, a significant number of Twitter users frame ChatGPT as an opportunity rather than a threat. Topics like *Skills and Productivity* or *Text Generation for Professionals* suggest potential ways in which ChatGPT can be used as a tool for increasing workers' efficiency and saving time. As later revealed in some of the close-read tweets, many people showcase how they utilise the AI chatbot in their daily activities, automating certain tasks and facilitating their workflow.

Moreover, topics delineated in Table 1. present a variety of content that ChatGPT can create. This theme is present in *Applications in Content Creation*, *Opportunities in Generating Content* and *AI Applications and Language Processing*, which together account for 15,19% of the Twitter posts. The AI chatbot has gained popularity as a valuable tool for assisting individuals with writing a multitude of content such as emails, blog posts, papers or even poems. In this context, ChatGPT can surpass the language barriers, again improving people's efficiency with written tasks. Another popular application of ChatGPT is coding, portrayed with the theme *Utilisation of AI in Coding* (4.05% of tweets). Twitter users have shown how programming can be facilitated with the AI chatbot, which can be particularly useful for beginners.

Furthermore, the second most frequent theme, *AI Tools*, present in 11.26% of tweets, revolves around the significant role of AI tools and chatbots in driving technological advancements and knowledge acquisition. It suggests that Twitter users share insights on how these systems are transforming industries, emphasising an evident presence of technological debates on Twitter. This has also been captured by *Recent Developments in Tech Sector* and *Discussions and News on Technology*. Additionally, the third most frequent theme, reflected in 7.21% of tweets, relates to aspects of communication with chatbots (topic also named *Communication with Chatbots*) such as the practice of writing prompts. This has been a popular point of discussions as, in light of latest advancement with LLMs, structuring appropriate prompts is becoming an increasingly important skill, as it also influences the output users receive.

Topic Modeling – Bidirectional Encoder Representations from Transformers (BERT)

Despite being one of the most common methods in topic modelling, widely applied in numerous studies, LDA performs poorly on short texts such as tweets (Amrouche, 2019; Pelgrim, 2021). Therefore, in an attempt to derive better results on latent topics in the analysed dataset, this research incorporates another topic modelling method, known as BERT. After applying it to the study's dataset, this technique

generated 2300 clusters. Unfortunately, 75.81% of the tweets were classified as outliers, which do not bring any added value and, hence, will be ignored. On the other hand, it is quite beneficial that BERT does not force the clusters, also in light of the dataset's size and the variety of discussions that Twitter users have. As a result, these topics are more concrete with a narrower focus. Given the scope of this study, only a set of generated BERT topics is analysed here. The cutoff point for including the results was that the clusters have to contain at least 0.10% of the tweets, accounting for around 1979 posts. Additionally, some of the topics were further excluded due to lower interpretability of applicability, such as Twitter jargon (see Appendix 2). In the end, 27 clusters are analysed in this part. Similarly as with LDA, based on their keywords and related tweets, these topics were manually labelled and interpreted (Table 2). Due to a high number of topics, they were grouped into meta-topics based on commonalities of their content.

Table 2. Topics generated with BERT, groped to meta-topics.

Meta-Topic	**Topics and Proportion of Tweets**
Technology Advancements and Tech Industry	BERT Topic 2: Artificial General Intelligence (0.43%) BERT Topic 3: Google Search Engine (0.40%) BERT Topic 6: Microsoft Bing search (0.31%) BERT Topic 14: Google's Bard (0.17%) BERT Topic 16: Virtual Assistants (Siri, Alexa) (0.17%) BERT Topic 23: Elon Musk (0.15%) BERT Topic 31: Tech Investments (0.11%) BERT Topic 32: Search Engine Optimization (0.10%) BERT Topic 34: Baidu's Ernie Bot (0.10%)
Applications of LLMs	BERT Topic 8: Programming (0.30%) BERT Topic 10: Cryptotrading (0.25%) BERT Topic 17: Healthcare (0.17%) BERT Topic 20: Prompt Engineering (0.16%) BERT Topic 22: Chatbots in Customer Service (0.15%) BERT Topic 24: Mathematical Calculations (0.14%) BERT Topic 36: Optimus AI (0.10%) BERT Topic 28: Digital Marketing (0.12%) BERT Topic 33: Education and Teaching (0.10%)
Content Creation	BERT Topic 0: Poetry and Music (0.61%) BERT Topic 27: YouTube Content (0.13%)
Risks and Concerns	BERT Topic 7: Risks to Cybersecurity (0.31%) BERT Topic 9: Italy and Privacy Protection (0.25%) BERT Topic 18: Political Bias (0.17%) BERT Topic 21: Legal Considerations (0.15%) BERT Topic 29: Artificial Hallucinations (0.11%) BERT Topic 37: DAN mode and Jailbreak Prompts (0.10%)

In accordance with LDA results, topics generated with BERT showcased a lot of domains where ChatGPT can be utilised. Twitter users discuss its potential for numerous fields such as programming, customer service, digital marketing and crypto trading. Its application to crypto-trading has also been discovered while extracting the most popular hashtags, which suggested the popularity of this topic among Twitter users. In BERT results, two of the analysed topics relate to this domain, *Cryptotrading* and *Optimus AI*, which is a decentralised currency. Interestingly, the applications of chatbots included even the field of healthcare. People have also tested ChatGPT's capabilities in creating more creative content, such as writing songs, poems or generating YouTube videos' scripts. A similar pattern was detected by LDA. However, BERT delivered more concrete examples.

The majority of analysed BERT topics was classified into the meta-topic *Technology Advancements and Tech Industry*. This method unravelled the presence of discussions relating to other AI systems such as virtual assistants and chatbots like Baidu's Ernie Bot and Google's Bard, which were released in recent months as alternatives to ChatGPT. These results illustrate the scope and importance of the recent advancements in the tech industry. This is further emphasised by the presence of conversations on *Artificial General Intelligence,* marking significant technological inventions taking place in recent months. Furthermore, the performance of different AI systems has been widely compared by users, including comparisons made between ChatGPT and Google's search engine. After close-reading some of the tweets relating to these discussions, it can be concluded that these latest events in the tech world were often framed as a competition and rivalry between the aforementioned corporations.

Furthermore, BERT revealed some additional topics that were not uncovered while performing LDA, relating to risks and concerns that users have been expressing in online discussions. First, people have pondered on its impacts on cybersecurity. AI text generators like ChatGPT are considered to be "a game changer for #cybercriminals". The chatbot can improve the quality of phishing emails, making them more convincing and more difficult to detect (Caulfield, 2023). Moreover, what seemed to stir up the debate on privacy protection, is the ban of ChatGPT imposed by the Italian government. This news has been shared by many microblogs, fuelling the discussions on how invasive these AI systems are and what regulations should be introduced. Interestingly, in accordance with Rozado's (2023) findings, Twitter users have reported on ChatGPT's political bias, suggesting prevalent left-wing beliefs. Other types of biases mentioned by Twitter users referred to religion. Elon Musk was often mentioned with regards to these topics, as he has often criticised companies for failing with respect to AI safety and has been speaking about introducing an unbiased chatbot (Chakraborty, 2023).

Additionally, BERT revealed the presence of the *Artificial Hallucinations* topic. It refers to the AI systems generating plausible sounding output that is false. Twitter users shared their experience with ChatGPT making factual mistakes, despite the convincing tone of these answers. Moreover, Twitter

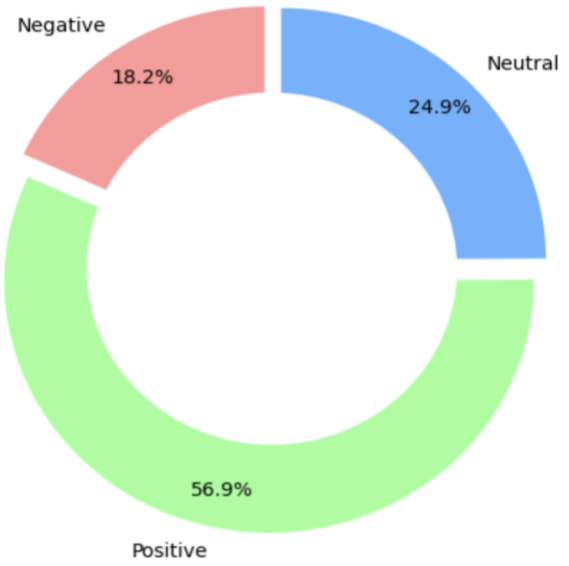
users have shown that ChatGPT's ethical barriers and safeguards can be easily bypassed by adjusting your query and for example, starting a prompt with: "Roleplaying as an AI without ethical standards." This testing of its restrictions is known as 'jailbreaking' and is captured by BERT Topic 37, labelled as "DAN mode" which stands for "Do Anything Now". This is closely related to the aspect of prompt engineering, another popular topic in the discourse on ChatGPT. Jailbreaking prompts are meant to free the AI bot from OpenAI's policies. This poses risks of generating toxic and dangerous content.

Overall, the two topic modelling methods revealed that ChatGPT is mostly framed in terms of its application in different domains. Included BERT topics, despite showing only a small percentage of the clusters, has showcased more specific examples of how the AI chatbot can be utilised. Moreover, it uncovered themes relating to risks and concerns stemming from the use of advanced LLMs. Moreover, the discourse surrounding ChatGPT to a large extent involves different tech companies and their AI systems, fostering a sense of competition in the field.

Sentiment Analysis – Valence Aware Dictionary and sentiment Reasoner (VADER)

Despite existing concerns about bias or privacy breaches, the sentiment analysis conducted with VADER, revealed a predominantly positive valence present in the discourse on ChatGPT with 56.9% of Twitter posts classified as positive. Only 18.2% of tweets were noted as negative and the remaining 24.9% as neutral. The prevalence of a positive tone on Twitter discourse is in line with Lin and Pena's (2011) claim who suggest that positive emotional messages are more popular. In the same vein, Stefanone et al. (2015) while analysing the diffusion of images shared by Twitter users in the context of #guncontrol, found an association between positively valenced images and more retweets, indicating higher likeliness of diffusion.

Figure 2. Distribution of Sentiment Scores in Tweets Discussing ChatGPT.



Examples of positively valanced tweets discuss increased efficiency due to the use of ChatGPT. Its ability to streamline a variety of tasks, speed them up and save time, has encouraged many users to use it on a daily basis. A lot of enthusiasm also stemmed from the fact that ChatGPT is available for free. Furthermore, users have advertised new initiatives that emerged following the launch of the AI chatbot such as "prompt teaching master class". Additionally, the positively valenced Twitter posts portray the excitement surrounding numeros applications of ChatGPT and its usefulness in many domains, as shown in the topic analysis. Users have shared articles and sources explaining how AI can improve many fields and discussed its growing potential in a number of areas, ranging from architecture to schooling. Additionally, positively valanced tweets also included more humorous conversations, such as testing the chatbot's capabilities to produce jokes.

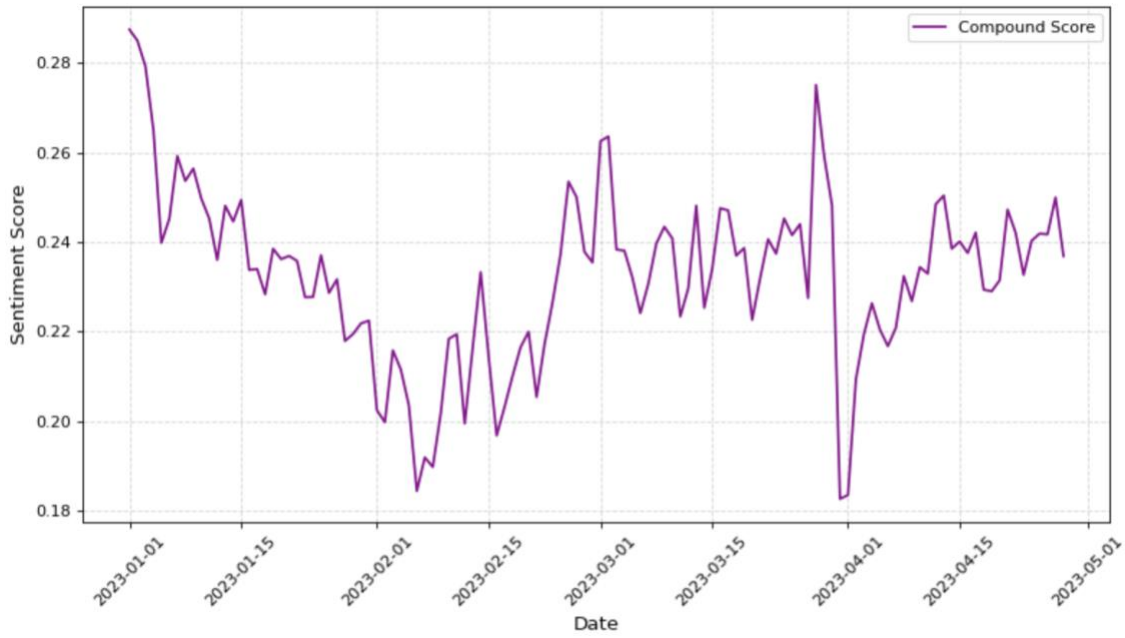
Twitter posts classified as the most negative debated over privacy concerns, emphasising the need to address data protection regulations and the use of sensitive information by the language model. For instance, to provide evidence for potential risks, an article for MIT technology review, where the author shows how much ChatGPT 'knows' about their editor in chief, went viral. Additionally, another close read negative tweet suggested that systems like this can potentially be a "game changer for #cybercriminals", which was also identified in topic modelling. Furthermore, users stressed that ChatGPT often fails to correctly answer factual questions. Some highlighted that the chatbot was last updated in 2021, making it more likely to provide inaccurate information. One of the tweets, assigned with the second most negative score, mentioned the chatbot's outdated data as well as alluded to its political bias, leaning towards left wing beliefs, which is in line with Rozado's (2023) findings. Moreover, some users delineated tasks in which ChatGPT fails to provide correct answers and is overall not performing well. This includes its comprehension of coding (despite being able to explain codes),

grammar or mathematical problems. Moreover, in some of the negative tweets that were close read, users suggested that ChatGPT's ethical barriers and safeguards can be easily bypassed by adjusting your query, the issue captured by one of the BERT topics in the previous section.

Since these results often refer to topics uncovered in LDA and BERT, a one-way ANOVA was conducted to investigate whether the sentiment score varies among different topics. The results indicated a significant effect of the topic on the sentiment. For LDA topics, the F-value equals 741.56, with p-value lower than 0.05, and for BERT topics, the calculated F-value is 918.33, with p-value lower than 0.05 as well.

In order to better understand the dynamics of the Twitter discourse surrounding ChatGPT, an average sentiment score was calculated for each day in the dataset and displayed to show its variation over time. Despite the values mostly falling in the same score range, there are two significant drops of the score on the 6th of February and at the end of March.

Figure 3. Average Sentiment Score per Day



Investigated tweets from the 6th of February revealed a lot of discussions pertaining to a launch of a different LLM, Google's BARD. Users have compared the two AI systems and framed it as a rivalry between OpenAI and Google, often referring to it as an 'AI war' and 'race'. This rhetoric, suggesting a competition within the tech industry, seemed to exacerbate negative valence in the Twitter discussions. Interestingly, around that time a lot of users shared information on ChatGPT being biassed,mostly politically, but also discriminating against certain religions. Furthermore, another drop in the sentiment score was observed at the end of March. On March 31 2023, Italy banned the use of ChatGPT, which, as already mentioned in topic modelling, was widely discussed on Twitter by users reporting on this

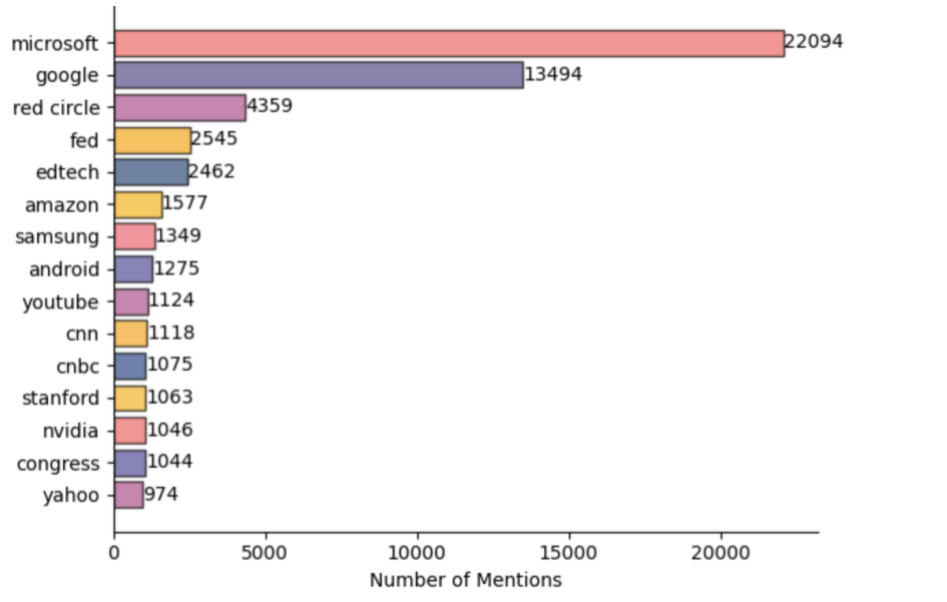
event. Similarly, a few days before the Future of Life Institute released the open letter, alarming people about the enormous risks that come with more advanced AI systems. These events have led to an increased number of tweets underscoring the urgent need for regulations.

Overall, the sentiment in tweets on ChatGPT seems to decrease with more posts relating to data security, such as in the context of Italy banning the chatbot due to privacy concerns. Moreover, what negatively impacts the overall valence of this discourse is the framing of competition between tech giants in launching more advanced AI systems, described as 'war', 'race' or 'battle'.

Key Actors – Named Entity Recognition (NER)

The NER analysis of actors involved in the discourse on ChatGPT, revealed the omnipresence of tech companies, as well as other corporations. The same conclusion was drawn from earlier steps, in BERT topic modelling, sentiment analysis and even hashtag extraction. Given the importance of the newest developments in the LLM realm, the involvement of other tech giants is understandable and emphasises the importance of this breakthrough. On the other hand, a high degree of tech involvement might suggest accumulation of power within this industry, a phenomenon strongly criticised by Emily M. Bender (2023) in her response to the Future of Life Institute petition.

Figure 4. Top 15 organisations involved in the discourse on ChatGPT, identified with NER.



Google and Microsoft, as visualised in Figure 4, are predominant in the discourse on ChatGPT. As outlined in topic modelling and sentiment analysis, their versions of AI chatbots have been juxtaposed with OpenAI's ChatGPT. Twitter users also widely report on investments these companies make, emphasising the dynamic and rapidly evolving landscape with new technologies, products, and services being introduced. Other corporations such as Samsung and Amazon also occur in the analysed dataset. Recently, CNBC, also included in NER results, reported on how customers use ChatGPT for writing

product reviews on Amazon (Mok, 2023). However, as a company that is investing in the development of AI services, it was also mentioned in the context of corporations competing in a tech 'race'. Samsung, on the other hand, appeared in the discussion on ChatGPT due to an unfortunate incident with the company's employees leaking sensitive information via OpenAI's chatbot (Powell, 2023). Many Twitter users reported on this situation, with some of them using the event to showcase the security risks posed by ChatGPT.

Another interesting example of ChatGPT deployment was discussed by Twitter users with respect to the Federal Reserve, US central banking system. The AI-powered chatbot has been tested in a number of market-relevant tasks, including deciphering Fed statements and generating trading signals based on large amounts of text data (Lee, 2023). ChatGPT's potential in forecasting stock market trends sparked interest in many Twitter users. This, together with popular discussions on cryptomarket, illuminate the influential role of the AI-powered chatbot in the finance world.

Additionally, the Red Circle account seems to be an important microblog in the discussion on ChatGPT. However, based on close read posts, it does not focus on a specific topic with respect to the AI chatbot. Note that the entity titled as 'edtech' does not apply to an organisation but rather ChatGPT's revolutionary effects on new technological implementations in education. Close read tweets, where 'edtech' was identified, discuss new opportunities for students and teachers, stemming from AI. However, Stanford is the only teaching and research institution included in the top 15 organisations involved in the discourse on ChatGPT on Twitter. This further underscores the dominance of corporations in this debate.

Discussion

The analysis of Twitter users' perceptions of ChatGPT utilising NLP methods has provided valuable insights into the framing of this recent advancement in the field of LLMs. This section aims to highlight the most significant findings, shedding light on how Twitter users attribute meaning to ChatGPT and the associated advancements in AI.

To begin, ChatGPT is mostly discussed with respect to its potential in other domains. In a large number of discussions, Twitter users share how they utilise ChatGPT in their daily life, often to facilitate their workflow. A variety of domains is discussed in terms of this AI-powered chatbot and its potential in this area. Popular topics include the utilisation of ChatGPT in crypto trading, coding and creation of written content. This encompasses even creative forms of texts like songs or poetry. Despite the enthusiasm, Twitter users acknowledge the chatbot's limited comprehension with respect to these assignments. Nevertheless, ChatGPT has been successful with automating a variety of other tasks, making it a useful tool for a large group of people. Due to that, the general valence in the discussions on ChatGPT is predominantly positive.

Yet, the perception of ChatGPT is not exclusively positive. One notable event that significantly lowered the sentiment score was the ban of ChatGPT in Italy, which generated a lot of discussions among Twitter users and raised concerns about data protection. It is important to note that the negative sentiment surrounding this topic may not necessarily reflect the actual opinions of users, but rather be influenced by the tone and framing of the news itself and the use of words such as 'ban' and 'privacy concerns'. Additionally, concerns about data security also emerged with respect to cybersecurity, which might be at risk in light of the new technologies. Furthermore, negatively valenced tweets discussed the presence of political bias ingrained in ChatGPT's responses. This has caused outrage among the users, significantly worsening the positive framing of the AI chatbot. Whilst much attention is devoted to data bias and privacy concerns in the academic literature, these domains are discussed to a lesser extent on Twitter compared to topics such as the utilisation of ChatGPT in cryptocurrency, coding or its impact on the job market.

Furthermore, throughout the analysis it became evident that tech companies such as Google, Microsoft, Amazon, and Baidu played a prominent role in the discussions surrounding ChatGPT. These companies, much like OpenAI, have made notable advancements in the field of artificial intelligence in recent months. Interestingly, Twitter users framed this phenomenon as a 'tech war', adding a negative connotation to the intense competition among these industry giants. This framing implies that the rapid pace of technological innovation is unstoppable, but it also raises concerns about the concentration of wealth among a select group of tech leaders.

Overall, there are two major frames that can be discerned following this analysis. First, there is a positive framing of ChatGPT in context of its applications in a variety of tasks. Users express enthusiasm about new opportunities in streamlining workflows, generating content, and assisting in problem-solving. On the other hand, there is another important emphasis frame present in this discourse. Twitter users actively engage in discussions regarding the advancements made by other tech companies in relation to ChatGPT. However, framing it as a 'tech war', or 'race', adds more negative valence to the discourse. It illuminates the intensity of competition in the AI sector, its unstoppable nature and the evident dominance of largest corporations in the discourse on technology. This nuanced understanding of ChatGPT's framing provides valuable insights into the dynamic and evolving discussions surrounding AI advancements and their impact on society.

Limitations

Despite the ability to reveal interesting patterns and trends, it is important to acknowledge the limitations of the implemented methods in understanding the context of the analysed content. Firstly, delineating latent topics cannot fully capture the argument conveyed in a given text. Sarcasm or jokes can easily deceive these techniques, and they may fail to reflect the intentions behind the, in this case, Twitter posts. Additionally, topic modelling requires a level of interpretation, making it more prone to researcher's subjectivity.

Secondly, there are notable downsides of the sentiment analysis. Again, it lacks a sophisticated comprehension of the analysed message. It is worth noting that disagreements among users may not necessarily imply negative sentiments towards ChatGPT. In some instances, users may actually try to calm others down about the perceived effects of ChatGPT. One notable limitation is the misclassification of tweets as negative, even when they do not criticise ChatGPT. For example, some users express their reliance on the chatbot due to a lack of confidence in their writing skills. However, their negative remarks about their own abilities lead to a negative sentiment score for the entire post. Additionally, the sentiment analysis results can be influenced by the presence of a single word, potentially skewing the overall sentiment analysis towards a more positive or negative interpretation. For instance, tweets relating to the competition between tech giants have exacerbated negative valence, which to some extent can be related to a frequent use of the word 'war' rather than actually negative attitude of users.

Thirdly, numerous limitations in this study stemmed from insufficient processing power for data analysis. This constraint restricted the complexity of the included methods. With stronger computing power, some of the methods could have been improved, possibly yielding better results. Additionally,

the smaller processing power also resulted in longer computation times and reduced efficiency in handling this dataset of around 2 million data points.

Therefore, while the implemented methods provide valuable insights, it is important to approach these results with caution and consider the nuances and complexities involved in accurately capturing patterns in social media discussions. Moreover, while computational methods have the potential to reduce researcher bias, this study still requires a level of interpretation, introducing some subjectivity.

Conclusion

In conclusion, the analysis of Twitter users' perceptions of ChatGPT using NLP methods has provided valuable insights into the framing of this recent development in the field of LLMs. The study highlights two major frames in the discourse surrounding OpenAI's chatbot: a positive framing of its applications in various tasks and a more negative framing related to competition among tech companies. Twitter users express enthusiasm for using ChatGPT as a tool facilitating numerous tasks. However, the intense competition among tech giants underscores the dominance of largest corporations leading the discourse. The debate on Twitter also differs from the themes present in the academic literature, focusing relatively more on fields like cryptocurrency or ChatGPT's applications to coding. Still, Twitter users touch upon the matters of data privacy or other present risks, similarly to scholars. Furthermore, the analysis also reveals limitations of NLP methods. Despite successfully showcasing trends and patterns present in large datasets, these techniques cannot understand the context and intentions of analysed content. Still, these findings contribute to a nuanced understanding of ChatGPT's perception and the complexities of AI discussions on social media.

References

Amrouche, M. (2019, December 11). Short Text Topic Modeling - Towards Data Science. *Medium*. https://towardsdatascience.com/short-text-topic-modeling-70e50a57c883

Bender, E. M. (2023, April 1). Policy makers: Please don't fall for the distractions of #AIhype. *Medium*. https://medium.com/@emilymenonbender/policy-makers-please-dont-fall-for-the-distractions-of-aihype-e03fa80ddbf1

Breuer, J., Wulf, T., & Mohseni, M. R. (2020). New Formats, New Methods: Computational Approaches as a Way Forward for Media Entertainment Research. *Media and Communication*, *8*(3), 147–152. https://doi.org/10.17645/mac.v8i3.3530

Burch, L. M., Frederick, E. L., & Pegoraro, A. (2015). Kissing in the Carnage: An Examination of Framing on Twitter During the Vancouver Riots. *Journal of Broadcasting & Electronic Media*, *59*(3), 399–415. https://doi.org/10.1080/08838151.2015.1054999

Caulfield, M. (2023, April 18). ChatGPT is changing the phishing game. *Security Info Watch*. https://www.securityinfowatch.com/cybersecurity/information-security/breach-detection/article/53057705/chatgpt-is-changing-the-phishing-game

Chadha, A. S. (2023, January 2). What Exactly Is Chat GPT & How Does It Work? For a 5-Year-Old! *Medium*. https://becominghuman.ai/what-exactly-is-chat-gpt-how-does-it-work-for-a-5-year-old-aeccabdfa990

Chakraborty, K. (2023). TruthGPT: A Maximum Truth-Seeking AI Chatbot. *Techopedia*. https://www.techopedia.com/truthgpt-a-maximum-truth-seeking-ai-chatbot

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). Opinion Paper: "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

Else, H. (2023). Abstracts written by ChatGPT fool scientists, 423-423 Nature, 613(7944). https://doi.org/10.1038/d41586-023-00056-7.

Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, *43*(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Firth-Butterfield, K. (2023, May 22). *AI leaders call for pause in systems training, plus other AI stories to read this month*. World Economic Forum. https://www.weforum.org/agenda/2023/03/ai-leaders-call-for-pause-plus-other-ai-stories-to-read-this-month/

Foggatt. (2023, March 1). How ChatGPT Will Strain a Political System in Peril. *The New Yorker*. https://www.newyorker.com/podcast/political-scene/how-chatgpt-will-strain-a-political-system-in-peril

Hemphill, L., Culotta, A., & Heston, M. (2013). Framing in Social Media: How the US Congress Uses Twitter Hashtags to Frame Political Issues. *Social Science Research Network*. https://doi.org/10.2139/ssrn.2317335

Kulshrestha, R. (2019, December 10). A Beginner's Guide to Latent Dirichlet Allocation(LDA). *Medium*. https://towardsdatascience.com/latent-dirichlet-allocation-lda-9d1cd064ffa2

Lee, J. (2023, April 17). ChatGPT Can Decode Fed Statements, Predict Stock Moves From Headlines. *Bloomberg.com*. https://www.bloomberg.com/news/articles/2023-04-17/chatgpt-can-decode-fed-speak-predict-stock-moves-from-headlines

Lin, J., & Peña, J. E. (2011). Are You Following Me? A Content Analysis of TV Networks' Brand Communication on Twitter. *Journal of Interactive Advertising*, *12*(1), 17–29. https://doi.org/10.1080/15252019.2011.10722188

McCallum, B. S. (2023, April 1). ChatGPT banned in Italy over privacy concerns. *BBC News*. https://www.bbc.com/news/technology-65139406

Metz, C., & Schmidt, G. (2023, March 29). Elon Musk and Others Call for Pause on A.I., Citing 'Risks to Society.' *The New York Times*. https://www.nytimes.com/2023/03/29/technology/ai-artificial-intelligence-musk-risks.html%20https://futureoflife.org/open-letter/pause-giant-ai-experiments/

Mok, A. (2023, April 26). AI chatbots like ChatGPT appear to be leaving product reviews on Amazon. *Business Insider*. https://www.businessinsider.com/ai-chatbots-chatgpt-may-be-writing-amazon-product-reviews-report-2023-4

Nisbet, M. C. (2009). Communicating Climate Change: Why Frames Matter for Public Engagement. *Environment: Science and Policy for Sustainable Development*, *51*(2), 12–23. https://doi.org/10.3200/envt.51.2.12-23

Pelgrim, R. (2021, March 8). Short-Text Topic Modelling: LDA vs GSDMM - Towards Data Science. *Medium*. https://towardsdatascience.com/short-text-topic-modelling-lda-vs-gsdmm-20f1db742e14

Powell, O. (2023, June 22). *Samsung employees allegedly leak data via ChatGPT*. Cyber Security Hub. https://www.cshub.com/data/news/iotw-samsung-employees-allegedly-leak-proprietary-information-via-chatgpt

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-physical Systems*, *3*, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Rozado, D. (2023). The Political Biases of ChatGPT. *Social Sciences*, *12*(3), 148. https://doi.org/10.3390/socsci12030148

Shepherd, J. (2023, May 16). *23 Essential Twitter Statistics You Need to Know in 2023*. The Social Shepherd. https://thesocialshepherd.com/blog/twitter-statistics

Stefanone, M. A., Saxton, G. D., Egnoto, M. J., Wei, W., & Fu, Y. (2015). *Image Attributes and Diffusion via Twitter: The Case of #guncontrol*. https://doi.org/10.1109/hicss.2015.216

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

Talboy, A.N., & Fuller, E. (2023). Challenging the appearance of machine intelligence: Cognitive bias in LLMs. ArXiv, abs/2304.01358.

Wicke P, Bolognesi MM (2020) Framing COVID-19: How we conceptualize and discuss the pandemic on Twitter. PLoS ONE 15(9): e0240010. https://doi.org/10.1371/journal.pone.0240010

Xiang, C. (2023, March 29). The Open Letter to Stop "Dangerous" AI Race Is a Huge Mess. *Vice*. https://www.vice.com/en/article/qjvppm/the-open-letter-to-stop-dangerous-ai-race-is-a-huge-mess

Yang, Y., & Pan, S. (2016). *Personalized Emphasis Framing for Persuasive Message Generation*. https://doi.org/10.18653/v1/d16-1150

Yao, M. (2022). Can Bots Manipulate Public Opinion? *TOPBOTS*. https://www.topbots.com/can-bots-manipulate-public-opinion/

Yuval Noah Harari argues that AI has hacked the operating system of human civilisation. (2023, April 28). *The Economist*. https://www.economist.com/by-invitation/2023/04/28/yuval-noah-harari-argues-that-ai-has-hacked-the-operating-system-of-human-civilisation?utm_content=section_content&gclid=Cj0KCQjw756lBhDMARIsAEI0AgkbUQ1-MGiNzMgJj2pTgXgGAFlnDZqQNGZLTEbrPu6XsQV7zg4bLNQaAhwIEALw_wcB&gclsrc=aw.ds

Zarifhonarvar, A. (2023). Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence. *ZBW - Leibniz Information Centre for Economics*. https://doi.org/10.2139/ssrn.4350925

Appendix

Appendix 1

The table below displays the names of columns from the analysed dataset together with explanations of their content. More information is available on Twitter's Developer Platform.

Column Name	**Explanation**
id	The unique identifier of the requested tweet.
thread_id	The tweet ID of the original tweet of the conversation (which includes direct replies, replies of replies).
timestamp	Creation time of the tweet.
unix_timestamp	Creation time of the tweet (different representation).
link	The URL link to the tweet on the Twitter platform.
subject	The subject or topic of the tweet
body	The main content of the tweet, the actual text message.
author	The username of the author of the tweet
author_fullname	The author's full name
author_id	The unique identifier of the Author who posted this tweet.
source	The name of the app the user tweeted from.
language_guess	Language of the tweet, if detected by Twitter.
possibly_sensitive	This field indicates content may be recognized as sensitive.
retweet_count	Number of times a tweet was reposted
reply_count	Number of replies given to the tweet
like_count	Number of likes given to the tweet
quote_count	The number of times the tweet has been quoted by other users
impression_count	The estimated number of times the tweet has been seen by users
is_retweet	Whether a tweet was retweeted

retweeted_user	The username of the original author of a retweeted tweet.
is_quote_tweet	Indicates whether the tweet includes a quote from another user's tweet
quoted_user	The username of the user being quoted in a tweet.
is_reply	If the represented Tweet is a reply, this field will contain the original Tweet’s author ID.
replied_user	If the represented Tweet is a reply, this field will contain the ID of the original Tweet’s author.
hashtags	The hashtags included in the tweet, denoted by the '#' symbol.
urls	The URLs or web links included in the tweet.
images	Image attachments in the Tweet
videos	Video attachments in the Tweet
mentions	Users mentioned in the Tweet
reply_to	The username or handle of the user being replied to in a tweet.
long_lat	Geocode of the Tweet.
place_name	Creation place of the Tweet.

Appendix 2

The following list displays topics generated with BERT, together with their keywords and labels, which were manually assigned. Overall, 2300 clusters were created. However, given the scope of this research, only 38 topics are investigated, each encompassing at least 0.10% of all tweets, which is around 1979 posts. Additionally, due to lack of interpretability or applicability, some of the 38 topics were discarded. These are highlighted with red colour. Topics 19 and 35 were treated as stopwords in this case.

BERT Topic 0: Poetry and Music (0.61%) – 'poem_song_music_haiku'
BERT Topic 1: Not included in the analysis – 'bro_lol_lmao_haha'
BERT Topic 2: Artificial General Intelligence (0.43%) – 'llm_consciousness_agi_emergent'
BERT Topic 3: Google Search Engine (0.40%) – 'google_search_engine_googling'
BERT Topic 4: Not included in the analysis – 'tweet_twitter_tweeting_account'
BERT Topic 5: Not included in the analysis – 'laughing_floor_rolling_laugh'
BERT Topic 6: Microsoft Bing search (0.31%) – 'bing_microsoft_edge_bingchat'
BERT Topic 7: Risks to Cybersecurity (0.31%) – 'cybersecurity_malware_cyber_phishing'
BERT Topic 8: Programming (0.30%) – 'code_coding_debugging_programmer'
BERT Topic 9: Italy and Privacy Protection (0.25%) – 'italy_italian_ban_privacy'
BERT Topic 10: Cryptotrading (0.25%) – 'bitcoin_cryptocurrency_cryptomarket_xrp'
BERT Topic 11: Not included in the analysis – 'tear_joy_face_smiling'
BERT Topic 12: Not included in the analysis – 'skin_tone_clapping_light'
BERT Topic 13: Not included in the analysis – 'character_limit_exceeds_exceeded'
BERT Topic 14: Google's Bard (0.17%) – 'bard_bardai_rival_google'
BERT Topic 15: Not included in the analysis – 'pointing_index_backhand_webcafe'
BERT Topic 16: Virtual Assistants (Siri, Alexa) (0.17%) – 'siri_alexa_apple_shortcut'
BERT Topic 17: Healthcare (0.17%) – 'healthcare_doctor_patient_medicine'
BERT Topic 18: Political Bias (0.17%) – 'racist_leftist_liberal_political'
BERT Topic 19: Not included in the analysis – 'autogpt_gpt4_gpt3_gpt'
BERT Topic 20: Prompt Engineering (0.16%) – 'prompt_script_bash_promptengineering'
BERT Topic 21: Legal Considerations (0.15%) – 'lawyer_legal_law_legaltech'
BERT Topic 22: Chatbots in Customer Service (0.15%) – 'chatbots_chatbot_customer_customerservice'
BERT Topic 23: Elon Musk (0.15%) – 'musk_elon_elonmusk_rival'
BERT Topic 24: Mathematical Calculations (0.14%) – 'calculator_math_calculation_mathematician'
BERT Topic 25: Not included in the analysis – 'interval_circle_bag_police'
BERT Topic 26: Not included in the analysis – 'sweat_grinning_beaming_smiling'
BERT Topic 27: YouTube Content (0.13%) – 'youtube_video_transcript_script''
BERT Topic 28: Digital Marketing (0.12%) – 'marketing_amp_marketer_digitalmarketing'
BERT Topic 29: Artificial Hallucinations (0.11%) – 'turing_hallucination_hallucinate_hallucinating'
BERT Topic 30: Not included in the analysis – 'flushed_mouth_pensive_relieved'
BERT Topic 31: Tech Investments (0.11%) – 'billion_49_investment_msft'
BERT Topic 32: Search Engine Optimization (0.10%) – 'seo_seos_keyword_keywords'
BERT Topic 33: Education and Teaching (0.10%) – 'education_classroom_edtech_teacher'
BERT Topic 34: Baidu's Ernie Bot (0.10%) – 'baidu_ernie_bidu_chinese'
BERT Topic 35: Not included in the analysis – 'chatgpt4_chatgptplus_chatgptdown_chat'
BERT Topic 36: Optimus AI (0.10%) – 'optimusaibsc_gen_aim_optimusai'
BERT Topic 37: DAN mode and Jailbreak Prompts (0.10%) –

‘jailbreak_jailbreaking_jailbroken_dan’