

Article

ChatGPT across Arabic Twitter: A Study of Topics, Sentiments, and Sarcasm

Shahad Al-Khalifa ¹, Fatima Alhumaidhi ¹, Hind Alotaibi ^{1,2}  and Hend S. Al-Khalifa ^{1,3,*} ¹ iWAN Research Group, King Saud University, Riyadh 11543, Saudi Arabia² College of Language Sciences, King Saud University, Riyadh 11421, Saudi Arabia³ Department of Information Technology, College of Computer and Information Sciences, King Saud University, P.O. Box 51178, Riyadh 11543, Saudi Arabia

* Correspondence: hendk@ksu.edu.sa

Abstract: While ChatGPT has gained global significance and widespread adoption, its exploration within specific cultural contexts, particularly within the Arab world, remains relatively limited. This study investigates the discussions among early Arab users in Arabic tweets related to ChatGPT, focusing on topics, sentiments, and the presence of sarcasm. Data analysis and topic-modeling techniques were employed to examine 34,760 Arabic tweets collected using specific keywords. This study revealed a strong interest within the Arabic-speaking community in ChatGPT technology, with prevalent discussions spanning various topics, including controversies, regional relevance, fake content, and sector-specific dialogues. Despite the enthusiasm, concerns regarding ethical risks and negative implications of ChatGPT's emergence were highlighted, indicating apprehension toward advanced artificial intelligence (AI) technology in language generation. Region-specific discussions underscored the diverse adoption of AI applications and ChatGPT technology. Sentiment analysis of the tweets demonstrated a predominantly neutral sentiment distribution (92.8%), suggesting a focus on objectivity and factuality over emotional expression. The prevalence of neutral sentiments indicated a preference for evidence-based reasoning and logical arguments, fostering constructive discussions influenced by cultural norms. Sarcasm was found in 4% of the tweets, distributed across various topics but not dominating the conversation. This study's implications include the need for AI developers to address ethical concerns and the importance of educating users about the technology's ethical considerations and risks. Policymakers should consider the regional relevance and potential scams, emphasizing the necessity for ethical guidelines and regulations.

Keywords: ChatGPT; Arabic tweets; topic modeling; sentiment analysis; sarcasm detection



Citation: Al-Khalifa, S.; Alhumaidhi, F.; Alotaibi, H.; Al-Khalifa, H.S. ChatGPT across Arabic Twitter: A Study of Topics, Sentiments, and Sarcasm. *Data* **2023**, *8*, 171. <https://doi.org/10.3390/data8110171>

Academic Editors: Hiram Calvo and Han Woo Park

Received: 11 September 2023

Revised: 24 October 2023

Accepted: 10 November 2023

Published: 14 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

ChatGPT, developed by OpenAI, is a language prediction model that has seen widespread use across various sectors globally. Since its inception in November 2022, the usage of ChatGPT has significantly expanded worldwide, with a plethora of use cases. According to [1], ChatGPT has garnered significant attention due to its remarkable ability to emulate human-like conversation and provide insightful responses to intricate inquiries. Through its training on vast amounts of text data, ChatGPT has acquired a deep understanding of various topics and can engage in conversations with users in a coherent and contextually appropriate manner. One of the most impressive aspects of ChatGPT is its capacity to handle complex questions. It can comprehend nuanced queries, break them down into meaningful parts, and generate well-formed responses that take into account the context and nuances of the conversation [2]. This capability has led to its widespread use in a variety of domains, including healthcare [3,4], business [5], education [6,7], and content generation [8].

However, it is important to note that ChatGPT is not without limitations. While it can produce impressive results, it is still prone to generating incorrect or nonsensical answers,

especially when faced with ambiguous or misleading questions [9]. Additionally, it may exhibit biases present in the training data, potentially leading to biased or discriminatory outputs [10,11]. Efforts are being made by developers and the research community to minimize these issues and improve the overall performance and reliability of the model.

Looking ahead, there are high expectations for the future development of ChatGPT and similar language models. OpenAI and other organizations are actively working on refining and enhancing these models to address their limitations. This includes refining the training methodologies, improving the model's ability to handle ambiguous queries, and implementing mechanisms to reduce biases [1,8,11]. According to [12,13], OpenAI has plans to release more powerful and capable versions of ChatGPT, which are expected to exhibit even more human-like conversational abilities. These advancements will likely involve training on larger datasets, incorporating more diverse sources of information, and fine-tuning the models based on user feedback and real-world usage.

As ChatGPT continues to evolve, it holds the potential to revolutionize various industries and applications. From personalized virtual assistants to intelligent tutoring systems and creative content generation, the further development of ChatGPT and similar language models are poised to have a profound impact on how we interact with and utilize AI-powered conversational systems in the years to come [14].

On the other hand, despite the global significance and rising adoption of this technology, there is a noticeable vacuum in its exploration within certain cultural contexts, specifically within the Arab world. This study explores the primary topics discussed by early Arab users in Arabic tweets that are related to ChatGPT using various topic modeling techniques. Additionally, this study aims to analyze the sentiments expressed by early Arab users in their ChatGPT-related tweets and to determine whether these tweets convey a serious or sarcastic tone.

To understand the perceptions of ChatGPT's early adopters, several studies have analyzed Twitter data to explore the sentiment and topics discussed related to ChatGPT. The results revealed that the overall sentiment of ChatGPT users can be described as predominantly positive [15–17]. Early users expressed overwhelmingly positive sentiments, finding the experience successful and appreciating its capabilities in various domains. Topics commonly discussed in relation to ChatGPT include disruptions to software development, entertainment, creativity, artificial intelligence (AI), search engines, education, writing, and question-answering [17–19]. However, alongside the positive sentiment, some users also expressed concerns about the potential ethical and legal implications of ChatGPT. These concerns revolved around issues such as the misuse of technology and its impact on educational aspects. Trust, transparency, and ethical considerations were identified as important factors influencing public perception of emerging technologies like ChatGPT. It is worth noting that some studies highlighted potential issues with ChatGPT, such as generating inaccurate results with high confidence and susceptibility to bias and discrimination inherent in language models. There were also reports of concerns about the legal and ethical consequences of using ChatGPT, e.g., [20,21]. However, as users started to utilize it and realize its benefits, apprehension was generally reduced [18].

While most of these results focused on studies conducted among English-speaking users, the lack of research on how Arabic-speaking users perceive ChatGPT shows a significant research gap that needs to be addressed. It is equally important to identify the prevalent topics, sentiments, and tones, such as seriousness or sarcasm, expressed by Arab users in their interaction with ChatGPT. This knowledge will provide valuable insights into the public perception of ChatGPT among Arabic-speaking users and help improve the platform's performance for this language community.

Considering the above, this study aims to answer the following research questions:

1. What are the primary topics discussed by early Arab users in Arabic tweets related to ChatGPT? This research question aims to identify the main themes or subjects that Arab users engage with when discussing ChatGPT on Twitter.

2. What sentiments are expressed by early Arab users in their ChatGPT-related tweets? This involves exploring the emotions expressed by Arab users in their tweets related to ChatGPT. It aims to uncover the prevailing sentiments associated with their interactions.
3. Do these tweets predominantly convey a serious or sarcastic tone? This research question focuses on determining the predominant tone used by Arab users in their tweets about ChatGPT. It investigates whether the overall tone is serious or sarcastic, providing insights into the communication style employed by Arab users.

Therefore, the purpose of this study is to explore these research questions using a variety of data analysis and topic modeling techniques.

The rest of the paper is organized as follows: First, we shed light on previous studies focusing on various aspects related to ChatGPT and public perceptions of the platform. Next, we present the methodology employed in the study, including data collection, data preprocessing, and analysis techniques. We then delve into the findings of the study, discussing Arab users' perceptions of ChatGPT, the common themes that emerge from the data, and the sentiment around ChatGPT in the Arab Twitter user community. Finally, we draw conclusions from the study and point toward future research directions in this field.

2. Related Work

ChatGPT, developed by OpenAI, has garnered significant attention due to its impressive performance in natural language understanding and generation. The model's ability to engage in human-like conversations and answer complex queries has marked a significant advancement in the field of natural language processing (NLP). Since its release, various studies have been conducted to assess ChatGPT's capabilities and limitations [11,21–23]. Several researchers have highlighted ChatGPT's strengths in handling a wide range of topics and its ability to provide coherent and contextually relevant responses [24]. Studies acknowledged the potential of ChatGPT in various fields, such as healthcare [3,13,25], education [6,7,26], academic research and practice [27,28], military [11,29], and tourism [30,31].

Despite its remarkable capabilities, ChatGPT has notable limitations. It can produce plausible-sounding but incorrect or nonsensical answers. It may also exhibit sensitivity to input phrasing and generate biased or inappropriate content. The model sometimes responds to harmful instructions or exhibits political and controversial biases, raising concerns about its responsible use. Some of the limitations of ChatGPT and AI-generated content are outlined in Erfina (see Figure 1) [15]:

- Bias and Fairness: NLP models like ChatGPT can inadvertently perpetuate biases present in training data. Ensuring fairness and reducing bias in AI-generated content is an ongoing challenge [10,11].
- Privacy: The generation of human-like text by AI models can pose privacy risks when misused to create convincing impersonations or for phishing attacks [32].
- Misinformation and Disinformation: AI-generated text can contribute to the spread of misinformation and disinformation, making it vital to detect and mitigate false or misleading content [33,34].
- Manipulation: AI can be used to manipulate public opinion, influence elections, or deceive individuals through persuasive and targeted content [29,35].
- Cybersecurity threats: The interactive nature of chat-based models presents opportunities for malicious actors to exploit vulnerabilities. Adversarial attacks, data poisoning, and manipulation of the model's responses are potential cybersecurity risks that need to be addressed [36,37].
- Social threats: The use of ChatGPT and other AI-powered chatbots may lead to job displacement and loss of human interaction [11,22].

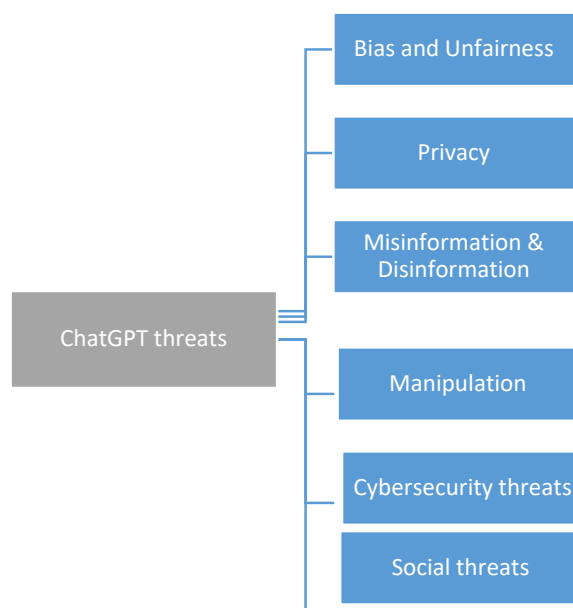


Figure 1. ChatGPT threats.

On the other hand, ChatGPT has undergone continuous advancements and updates [1]. OpenAI has made efforts to refine the model based on user feedback and address its limitations. These advancements aim to enhance the overall performance, reliability, and user experience of ChatGPT. Techniques such as fine-tuning and reinforcement learning have been employed to improve ChatGPT's responses and make them more aligned with user preferences.

According to [38], science, like other sectors of society, is currently concerned with the impact of AI technology, which challenges its core values, practices, and standards. However, rather than resisting the change, it is crucial to embrace the potential benefits while effectively managing the associated risks. Addressing ethical considerations and cybersecurity concerns is crucial for the responsible development and deployment of ChatGPT and similar models. Exploring users' experiences and perceptions when using such technologies can play a key role in improving the performance of ChatGPT and addressing ethical and cybersecurity issues. These studies can enhance the research and development of these tools to ensure that they are beneficial, fair, and secure for users and society as a whole. In the next section, we explore studies on ChatGPT perceptions among its users.

2.1. Sentiments of ChatGPT's Users

Since its release, ChatGPT has attracted significant attention from researchers and the public alike, with many studies being conducted to investigate public perceptions of the platform. While some studies discussed the broad social impact of ChatGPT [11,13,39,40], other studies have explored more specific aspects, including user experiences [41,42], attitudes [20,43], and concerns around its ethical use [6,22,44]. Several studies focused on understanding the sentiments of ChatGPT's early adopters, hoping to provide insights into its strengths and weaknesses from users' perspectives [16,17,20,45,46]. According to Birjali et al. [19], sentiment analysis refers to the task of extracting and analyzing people's opinions, sentiments, attitudes, perceptions, etc., toward different entities such as topics, products, and services. Early adopters of a product usually play a critical role in shaping the success or failure of new technology [17]. Their enthusiasm and influence can help to generate buzz and interest in a product, and their feedback can provide valuable insights into areas for improvement. As such, understanding the opinions and sentiments of early adopters is key to predicting the potential success or failure of a product in the marketplace. Twitter has been a popular social media platform used in many studies

to investigate the sentiment of ChatGPT users. Tweets are considered a representative sample of public attitudes towards ChatGPT due to their accessibility, real-time nature, and suitability for analysis using natural language processing techniques. The majority of early ChatGPT users' tweets expressed overwhelmingly positive sentiments related to topics such as disruptions to software development, entertainment, and exercising creativity. However, negative emotions such as fear and concern were also observed in some tweets. In Erfina [45], for instance, the sentiments of Twitter users towards ChatGPT were investigated, where a sample of 5000 English tweets was analyzed. The study revealed that the majority of Twitter users had a positive sentiment (57.6%) towards ChatGPT, while the negative sentiment reached 42.4%. While that study provides valuable insights into user sentiment towards ChatGPT, it is important to consider the limited sample size, i.e., 5000 tweets, which might negatively affect the generalizability of results. With a larger dataset but similar results, Korkmaz et al. [16] collected and analyzed 787,886 English tweets. Findings showed that despite encountering negative words frequently in tweets about ChatGPT, users generally perceive their experience as positive. The most frequently used words in tweets were associated with positive emotions, which contributed to the positive sentiment of the tweets. Even when tweets contained negative words, they were used with relatively less intensity, which contributed to the overall positive sentiment. Sharma et al. [46], on the other hand, reported different findings. They found that the majority of tweets related to ChatGPT are neutral, with a smaller proportion being positive or negative. The researchers also identified specific words and phrases associated with positive and negative sentiments. For example, words like "AI" and "language model" were associated with positive sentiment, while words like "bias" and "privacy" were associated with negative sentiment. The study suggests that public perception of emerging technologies like ChatGPT is influenced by trust, transparency, and ethical considerations. Although some studies focused only on topic modeling when analyzing tweets about ChatGPT (e.g., [47]), several studies adopted a mixed-method approach incorporating sentiment analysis and topic modeling to gain deeper insights. A description of these studies is presented next.

2.2. Topic Analysis of ChatGPT-Related Tweets

Topic modeling is a statistical technique used in natural language processing (NLP) to discover the abstract "topics" that occur in a corpus. Some of the most common algorithms for topic modeling include Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF) [48]. Researchers often combine sentiment analysis and topic modeling techniques to obtain a comprehensive understanding of the data. In a study by Haque et al. [17], for example, the researchers looked at 10,732 English tweets from early ChatGPT users. The researchers first identified the main topics using topic modeling and then performed a qualitative sentiment analysis for each topic. The results indicated that the majority of early adopters expressed overwhelmingly positive sentiments regarding topics such as disruptions to software development, entertainment, and creativity. However, some users expressed concerns regarding issues such as the potential misuse of ChatGPT, particularly with its impact on educational aspects. As the dataset used in this study was relatively small and collected over a shorter period of time, it does not offer a comprehensive understanding of the overall public attitude towards ChatGPT. In Tounsi et al. [18], there were reports of concerns about the potential ethical and legal implications of ChatGPT. A total of 463,983 English tweets were analyzed, and the results indicated that most users were satisfied with the tool, but some expressed worries about its possible legal and ethical consequences. Nevertheless, as people started to use it and realize its benefits for personal and business purposes, their apprehension was largely reduced. The article concludes by emphasizing the significance of acknowledging any potential downsides of new technology and utilizing it responsibly. Similar findings were reported by Koonchanok et al. [20]. The researchers analyzed 174,840 English tweets, and the results show that the overall sentiment towards ChatGPT is largely neutral to positive, and the most popular topics discussed

in tweets are AI, Search Engines, Education, Writing, and Question Answering. Their study also highlights potential issues with ChatGPT, such as its tendency to generate inaccurate results with high confidence and its susceptibility to bias and discrimination inherent in LLMs from their training datasets. With a larger dataset, including more than 150 scientific papers and 330,000 English and non-English tweets, Leiter et al. [21] conducted a more comprehensive analysis, classifying the dataset into 19 predetermined topics using a RoBERTa-based model. Their study found that ChatGPT is generally viewed positively, with emotions of joy dominating on social media, and it is characterized as a great opportunity across various fields in scientific papers, including the medical domain, but also as a threat to ethics and with mixed assessments for education. The study also noted a slight decrease in the perception of ChatGPT since its debut, with joy decreasing and negative surprise increasing, particularly in languages other than English.

After reviewing the existing literature, it is evident that there is a significant research gap regarding the perceptions of Arabic-speaking users toward ChatGPT despite its growing popularity within this demographic. To address this gap, it is essential to conduct studies that analyze Arabic tweets to gain valuable insights into the public perception of ChatGPT among Arabic-speaking users. Such studies could provide valuable feedback to improve the platform's performance for this language community and ensure that it meets their needs and expectations. Thus, the objective of this study is to explore the primary topics discussed on Twitter by early ChatGPT Arabic-speaking users. Additionally, the study aims to analyze the sentiments expressed by early Arab users in their ChatGPT-related tweets and to determine whether these tweets convey a serious or sarcastic tone.

3. Methodology

The methodology section outlines the research process, as depicted in Figure 2.

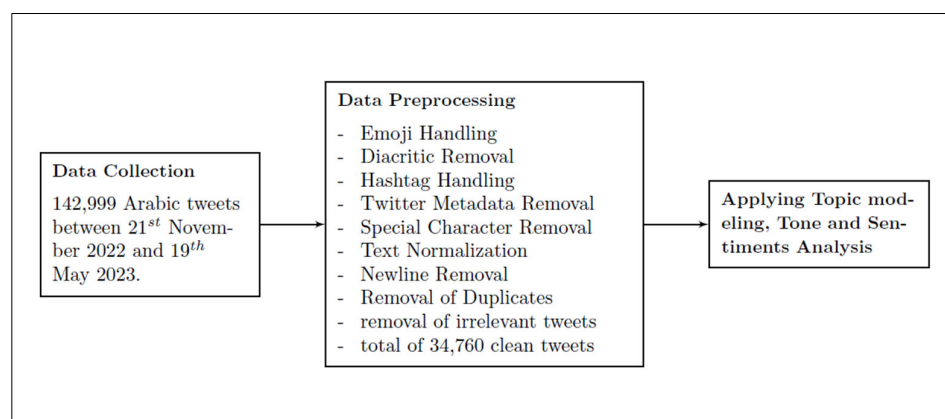


Figure 2. Flowchart of the research process.

The initial step involved data collection using two Python libraries: `snsrape` and `tweepy` [48]. The latter utilized the academic Twitter API. The main objective was to gather Arabic tweets discussing ChatGPT. Variations of the term “ChatGPT” in Arabic, such as “تشات جي بي تي” “تشات جيت” and “تشات جيت”, were used as keywords during data-collection phase, spanning from 21 November 2022 to 19 May 2023.

3.1. Data Processing

Arabic social media texts often contain noise in the form of elongations, diacritical marks, and extra symbols. Thus, data cleaning is crucial, involving the removal of punctuation, links, special characters, diacritics, emojis, and new lines. Additionally, certain Arabic letters can have different forms based on their position in a word; these letters were normalized for data simplification. The dataset was further refined by eliminating duplicates, irrelevant tweets, and very short ones, resulting in a final dataset of 34,760 tweets.

The preprocessing steps included the following:

- **Diacritic Removal:** Diacritics were eliminated from the text using the `dediac_ar` function from the `camel-tools` library to simplify the text and reduce linguistic variations.
- **Hashtag Handling:** Hashtags were processed by separating them into individual words and replacing underscores with spaces using the `handle-hashtags` function, ensuring they were distinct entities during topic modeling.
- **Emoji Removal:** Emojis were removed from the tweets using the `demoji` library for standardization, eliminating potential emoji influence on topic modeling.
- **Twitter Metadata Removal:** Twitter-specific metadata, such as usernames and links, were removed using regular expressions, eliminating noise, and focusing solely on tweet content.
- **Special Character Removal:** Special characters were stripped from the text to eliminate unnecessary symbols that could affect topic modeling.
- **Text Normalization:** Arabic words were normalized using techniques such as normalizing *alef*, *teh marbuta*, and *alef maksura* through functions from the `Camel-tools` library, simplifying vocabulary and aiding dictionary construction.
- **Newline Removal:** Newline characters were removed to maintain consistent formatting and eliminate unnecessary line breaks within tweets.
- **Duplicate Removal:** Duplicate tweets were removed to ensure uniqueness and avoid redundancy.
- **Manual Removal of Irrelevant Tweets:** About 600 irrelevant tweets were manually removed, addressing the issue of irrelevant documents in the corpus caused by Twitter's search.

The choice of 34,760 tweets as the dataset size was determined by several factors, which need to be clarified to address any potential concerns regarding its adequacy.

Firstly, Twitter's API imposes limitations on the amount of data that can be collected, both in terms of response size and the number of requests allowed per day. These restrictions necessitated careful management of our data-collection efforts. We adhered to the API limitations to ensure compliance and avoid any potential issues.

Furthermore, during the data-collection process, we encountered challenges related to data quality. It became evident that a significant portion of the collected tweets were either duplicates or too short to provide meaningful insights. Eliminating these duplicates and very short tweets was necessary to ensure the integrity and reliability of our dataset. By filtering out redundant and insufficient tweets, we aimed to focus on the most relevant and informative content for our analysis.

The selected dataset size strikes a balance between the constraints imposed by Twitter's API, the need for data quality, and the resources available to our research team. While we acknowledge that the dataset size may not be exhaustive, it allows for a substantial and diverse sample of Arabic social media content, which is crucial for our analysis. Despite the inherent limitations, we believe that the dataset provides valuable insights into the perceptions and interactions of Arab users with ChatGPT on Twitter.

3.2. Topic Modeling

To address the first research question regarding the primary topics discussed by early Arab users in Arabic tweets related to ChatGPT, several topic-modeling algorithms were employed (namely LDA and NMF); however, BERTopic was chosen for its outstanding performance on Arabic documents [42]. BERTopic combines transformer-based BERT embedding and class-based TF-IDF to generate coherent and interpretable topics. Documents are clustered based on semantic similarity using BERT embeddings, and class-based TF-IDF ranks word importance within topics [49].

BERTopic has been adopted in various text-processing tasks. Prior studies, e.g., [49,50], have showcased BERTopic's superiority in extracting meaningful topics from diverse Arabic textual datasets as opposed to other topic-modeling algorithms.

The BERTopic model was initialized with an Arabic BERT-based embedding model (`aubmindlab/bert-base-arabertv02-twitter`) and a representation model (`gpt-3.5-turbo`) from

OpenAI. The model was trained on the dataset, specifying 50 topics using trial and error. Unlike other algorithms, traditional coherence measures such as the C_v coherence score cannot be applied directly to BERTopic due to the representational model's encoding of text semantics and context. As a result, evaluating the coherence of topics generated by BERTopic using traditional coherence measures may not be applicable or meaningful. Therefore, the quality and interpretability of BERTopic-generated topics are assessed through qualitative analysis. Topics are inspected manually by examining sentences associated with each topic, evaluating their relevance to topic documents, and assessing coherence and consistency.

3.3. Sentiment Analysis

After conducting topic modeling, we performed sentiment analysis on the Arabic tweets with the classified topics to gain a deeper understanding of the emotional context within different thematic discussions on Twitter. The sentiment analysis was carried out with an unsupervised approach using the ArabiTools sentiment analysis pre-trained model available on Hugging Face (<https://huggingface.co/spaces/asalhi85/ArabiToolsSentimentAnalysis>, accessed on 10 November 2023), which was trained on a large-scale dataset of Arabic tweets. The tweets were classified as positive, negative, or neutral. Table 1 displays examples of the classified tweets.

Table 1. Examples of positive, neutral, and negative tweets.

Positive	Neutral	Negative
شات جي بي تي خورافي لكتابه الاميلات "ChatGPT is amazing for writing emails"	مؤسس تشات جي بي تي يرغب في افتتاح مكتب باليابان "The founder of ChatGBT wants to open an office in Japan"	اللي سوا chatgpt ما سوا فينا خير "The one who made ChatGPT did us no good"
شات جبت تاج راس الجميع "ChatGPT is a crown on everybody's head"	هل سيستطيع تشات جي بي تي الحفاظ علي انتشاره العالمي؟ "Can GBT chat maintain its global spread?"	chatgpt اكبر منافق عرفه التاريخ "ChatGPT is the biggest hypocrite in history"

3.4. Sarcasm Detection

In addition to performing sentiment analysis on the Arabic tweets with classified topics, exploring the dataset revealed the existence of sarcasm within the sentiment-labeled groups, especially in both the negative and positive sentiments (an example is shown in Table 2 below).

Table 2. Examples of sarcasm in positive and negative sentiments.

Positive	Negative
ههههه باين علي الشايب جي بي تي شارب شيء مخلل او مفلغل "Haha, it seems like the old GPT has had some pickled or spicy stuff."	والله أنا مش خايف علي المعلم شات جي بي تي غير مننا شويه وهايطلع يجري ورانا بالطوب "Swear to God, I'm not worried about the master ChatGPT except from us. Eventually, it will chase us with bricks."
يلا ممتاز لما نتخرج بكون chatgpt بعمل كلشي وبلا منا ومن شهادتنا هههه "Great, by the time we graduate, ChatGPT will be doing everything with no need to us or our degrees half haha."	انا لقيت حد منزل كورس عن ازاي تستخدم ال chatgpt يعيش النصاب علي قفي العبيط "I found someone who has uploaded a course on how to use ChatGPT. A scammer thrives where fools abound."

This finding encouraged us to conduct sarcasm detection, realizing its importance in comprehending the emotional expressions present in the dataset. Detecting sarcasm within tweets is essential because sarcastic statements can be wrongly classified by sentiment

analysis models due to their literal wording. Sarcasm detection was carried out with an unsupervised approach using the MARBERT Sarcasm Detector pre-trained model, which was fine-tuned on the ArSarcasT corpus (<https://huggingface.co/MohamedGalal/marbert-sarcasm-detector>, accessed on 10 November 2023). We utilized MARBERT for sarcasm detection in our analysis, drawing upon its demonstrated efficiency and robust performance from prior research, which underscored its ability to discern sarcasm with notable accuracy and reliability compared to other models [51].

A detailed discussion of the results is presented next.

4. Results and Discussion

4.1. Topic Modeling

Table 3 below presents the representation of the 50 BERTopic topics, including the count and ratio of assigned documents.

Table 3. BERTopic topic representation.

Topic	Topic Description	Count	Ratio
−1	ChatGPT—AI-powered chatbot developed by OpenAI	7482	0.215247
0	Using ChatGPT in AI-powered chatbots and search engines	16,979	0.488464
1	Requesting help in creating a ChatGPT account in Egypt	4202	0.120886
2	The impact of ChatGPT and AI on tech and people	2122	0.061047
3	Controversy in Colombia over use of AI-powered chatbot by judge to issue ruling	632	0.018182
4	The potential ethical risks of artificial intelligence (AI) and chatbots (ChatGPT) on human society	555	0.015967
5	Artificial intelligence and football—Elon Musk’s research on ChatGPT to develop a new alternative to chatbot technology for his company	337	0.009695
6	Cryptocurrency market analysis and predictions using ChatGPT and bitcoin as key components	239	0.006876
7	ChatGPT ban in Italy due to privacy concerns and illegal collection of personal data	153	0.004402
8	Alibaba to launch AI chatbot Tongyi Qianwen to compete with ChatGPT during cloud services conference next Tuesday	149	0.004287
9	AI tools for content creation and writing: ChatGPT, tweetmonk, aiseo art, blackbox, and beautifulai	133	0.003826
10	Ban on using chatbot AI tool by students at Sciences Po Institute in Paris	129	0.003711
11	Enhancing customer service using ChatGPT	103	0.002963
12	How to register and use ChatGPT with an American phone number or VPN in Saudi Arabia	100	0.002877
13	The impact of artificial intelligence and ChatGPT on skill development and government sectors at WGS2023 in the UAE	90	0.002589
14	The dark side of artificial intelligence-generated texts and images	85	0.002445
15	Educational events in Al-Batinah, Oman, featuring ChatGPT and artificial intelligence	82	0.002359
16	Using ChatGPT in Libya with troubleshooting tips	81	0.002330
17	ChatGPT—an AI-powered chatbot for generating human-like text responses using pre-trained transformer models on large language datasets	80	0.002301
18	Concluding Saudi ChatGPT Hackathon with over 650 participants from Academic Institution of Taif	73	0.002100
19	Chatbots and artificial intelligence in Morocco: challenges and opportunities	68	0.001956
20	Natural resources and future of pharmacy in Sudan	65	0.001870
21	The future of politics and economy in Kuwait with AI, education reform suggestions, and chatbot technology	55	0.001582
22	ChatGPT—An impressive AI application, but limited availability in Arab countries and requires a workaround for installation	54	0.001554
23	ChatGPT banned in Italy due to lack of privacy respect and absence of a system to verify the age of underage users	54	0.001554
24	Impact of AI and chatbots on marketing and society in Morocco	53	0.001525
25	Exploring the controversy surrounding ChatGPT in Morocco	51	0.001467
26	Discussion on ChatGPT and its role in Libya with concerns about hate speech and fake news verification	49	0.001410
27	Workshop on ChatGPT basics and employment opportunities in the education sector under the Mohammed bin Zayed prize for the best teacher	38	0.001093

Table 3. Cont.

Topic	Topic Description	Count	Ratio
28	The impact of OpenAI's GPT-3 on human jobs and governance	35	0.001007
29	Saudi Arabian football clubs and their achievements in Asian championships	30	0.000863
30	Artificial intelligence assistance in writing academic articles at Cardiff University and MIT	29	0.000834
31	Reviewing the positive and negative aspects of ChatGPT technology by Saudi Citizens living in the USA	29	0.000834
32	ChatGPT-3 discussing the Egyptian government's efforts in developing Sinai with officials and media outlets	28	0.000806
33	Films and storytelling	28	0.000806
34	OpenAI chatbot assistance for complex Java class and other technical requirements	26	0.000748
35	Cultural and genetic diversity of African and Arab tribes and clans	25	0.000719
36	Soccer clubs in Egypt—Al Ahly vs. Zamalek	25	0.000719
37	ChatGPT Plus subscription plan benefits and availability	24	0.000690
38	Yemeni crisis and international devastation: Ahmed Al-Sana'i's perspectives on the conflict	24	0.000690
39	The use of artificial intelligence in the government sectors—study requested by Sheikh Mohammed bin Rashid	23	0.000662
40	Chatbots with AI for natural language processing and programming capabilities	21	0.000604
41	Eliza AI model pushes victim to suicide in Belgium	19	0.000547
42	Understanding ChatGPT and rule-based AI models	18	0.000518
43	Grapeswap and Grape Token for passive income and farming in Arabic	17	0.000489
44	Issues and advancements in ChatGPT technology	15	0.000432
45	The impacts of artificial intelligence and future technology on employment and mental health in the Arab world	14	0.000403
46	Yemeni media reporting on the detainment of ChatGPT and its impact on national unity and economic stability	13	0.000374
47	The impact of AI on various fields, including music, search engines, medical diagnosis, and chatbots	12	0.000345
48	Beware of fake ChatGPT apps and websites: tips to avoid scams and protect your data	12	0.000345

Hierarchical clustering of these topics is depicted in Figure 3 below.

It is essential to address the redundancy in the discovered topics. Data collection from social media platforms can result in a significant amount of redundant information as users often share or re-post similar content. The dominance of a single topic indicates that a large portion of the tweets in the corpus share similar content and are focused on repeating news and information about GPT technology. This redundancy can be attributed to the nature of social media, where trending topics and viral news often receive extensive attention and discussion. As a result, topic-modeling algorithms may prioritize these dominant topics, leading to the overshadowing of other potentially valuable themes and discussions within the corpus. BERTopic's representational model provided valuable insights into the different topics discovered by the model. The use of contextual embeddings from the GPT-3.5 Turbo model allowed BERTopic to capture the semantic context of the tweets, making the documents in the corpus much clearer, especially regarding the sentiment in the documents. BERTopic's ability to consider sentiment within its topic-modeling process may have provided additional context and enriched the understanding of the discussions surrounding ChatGPT.

BERTopic's representational model yields insights into the distinct discovered topics. Leveraging contextual embeddings from the GPT-3.5 Turbo model enhances BERTopic's ability to capture semantic context in tweets. This clarity extends to sentiment understanding in documents. By considering sentiment, BERTopic enriches comprehension of ChatGPT-related discussions. Additionally, BERTopic's hierarchical clustering allowed for a more comprehensive exploration of topic relationships. This visualization helped identify potential clusters or groups of related topics, offering a more coherent and holistic view of the overall thematic landscape within the dataset.

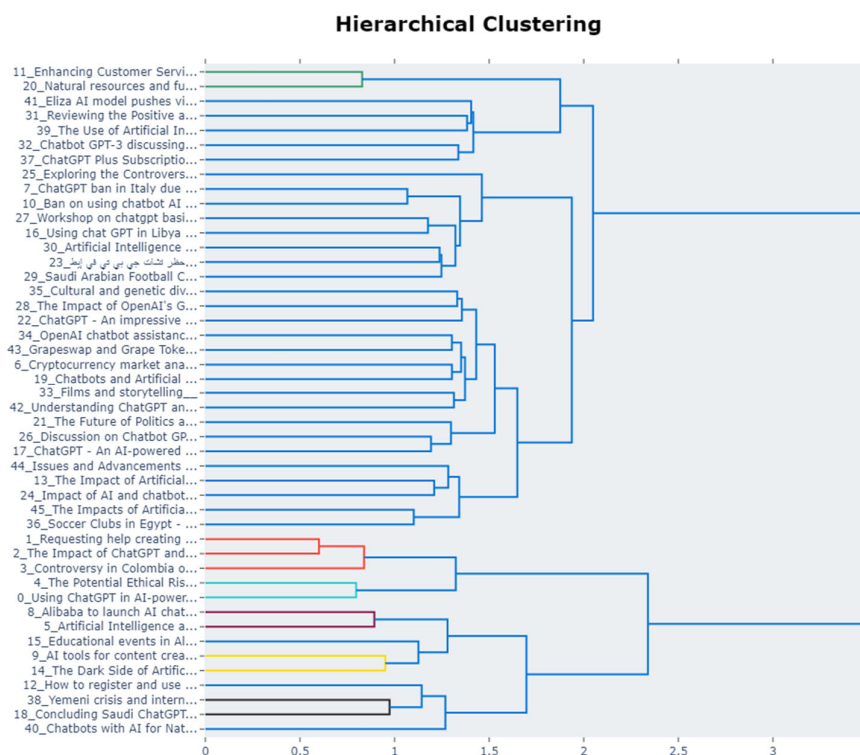


Figure 3. BERTopic hierarchical clustering, where the translation of the Arabic topic is “Abundance of ChatGPT”, and the red color indicates the most distinct cluster, and other colors represent secondary or less distinct clusters.

It is worth noting that certain topics might appear to be repeated due to the high diversity of dialects present in the dataset. The model might not always connect two tweets discussing the same topic, especially if they are in significantly different dialects. As a result, the distribution of tweets across topics can be seen as partially influenced by the variation in dialects. This phenomenon highlights the challenge of accurately clustering content in a multi-dialect dataset with diverse linguistic characteristics. Based on the identified topics, it becomes possible to establish connections between the discussed topics and events or trends from the preceding period. This contextualization sheds light on why certain topics underwent extensive discussions compared to others. Key insights and patterns from the results encompass the following:

1. **Dominant Topic:** The prevalence of tweets discussing ChatGPT technology signifies robust interest within the Arabic-speaking community. This dominance reflects ChatGPT’s popularity and its multidisciplinary influence, driving extensive social media discussions.
2. **Concerns and Controversies:** Topics covering controversies, ethical risks, and negative AI chatbot implications emerge. These discussions underscore apprehension around advanced AI technology, especially in language generation.
3. **Regional Relevance:** Region-specific AI applications and chatbot technology discussions in places like Saudi Arabia, Libya, Egypt, and Morocco underscore regional adoption and relevance in diverse contexts.
4. **Fake Content and Scams:** Topics surrounding fake ChatGPT apps and websites highlight concerns about fraud and potential AI-generated content scams.
5. **Sector-specific Dialogues:** Topics exploring AI and chatbot impacts on sectors like education, government, marketing, sports, and employment denote interest in understanding these technologies’ potential in various fields. These discussions reflect user interest in the benefits and challenges of AI integration across domains.

4.2. Sentiment Analysis

To address the second research question, we analyzed a total of 34,760 Arabic tweets classified into 50 BERTopic topics. The distribution of sentiments across all tweets resulted in 32,267 neutral tweets (92.8%), 1348 positive tweets (3.9%), and 1145 negative tweets (3.3%).

Our goal was to see the sentiment distribution for each topic and if there are some topics that have a notably positive, negative, or neutral sentiment. Figure 4 below shows the sentiment distribution for 49 topics, excluding –1, which are unknown topics.

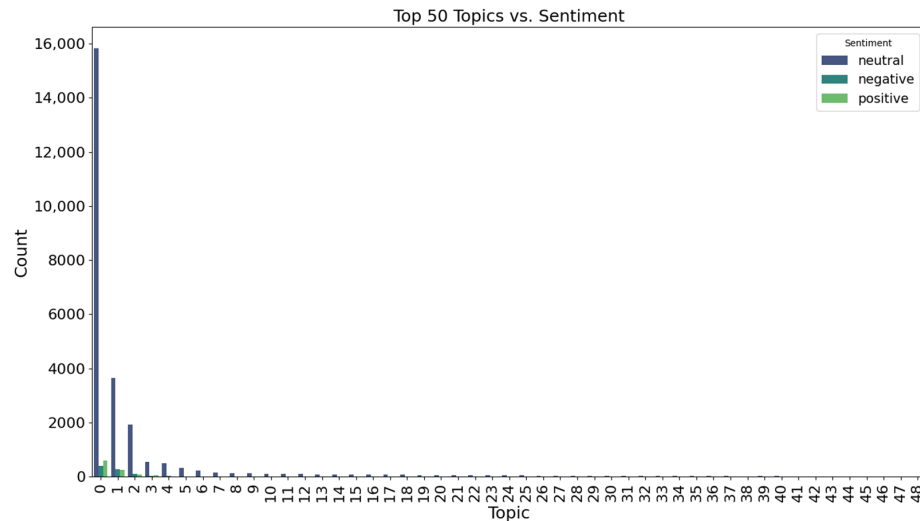


Figure 4. Distribution of sentiment on 50 topics.

To examine the distribution more closely, we visualized the distribution for the first 10 topics (see Figure 5).

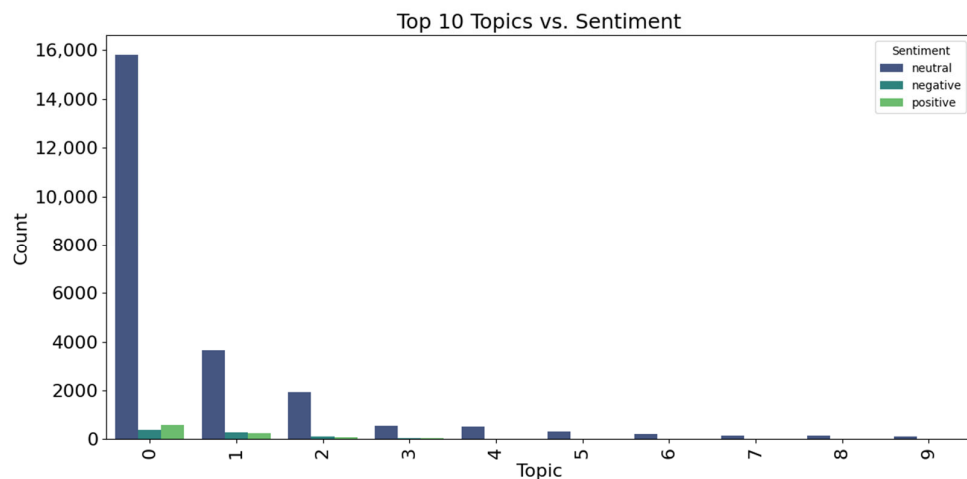


Figure 5. Distribution of sentiment on the top 10 topics.

We can see that the majority of sentiment in all topics is neutral, indicating that they tend to be objective or fact-based rather than being heavily influenced by personal feelings or biases. Moreover, Topic 0, with the description “Using ChatGPT in AI-powered chatbots and search engines”, has a higher proportion of positive sentiment in comparison to negative sentiment, unlike the other topics, which contain a higher proportion of negative sentiment in comparison to positive sentiment.

The variation in sentiment distributions across topics demonstrates how different thematic discussions on Twitter can result in different emotional responses from users.

It is worth noting that topics 0 and 1 had a high proportion of neutral sentiment tweets, likely because the majority of tweets in these topics were simply sharing news updates rather than expressing personal opinions. This indicates that users were largely tweeting in an informative manner to spread information rather than expressing their own reactions and opinions.

4.3. Sarcasm Detection

We were able to identify 1391 sarcastic tweets out of the 34,760 Arabic tweets (4%). The sarcastic tweets were spread across different topics, with some topics having a higher proportion of sarcasm than others (see Figure 6 below).

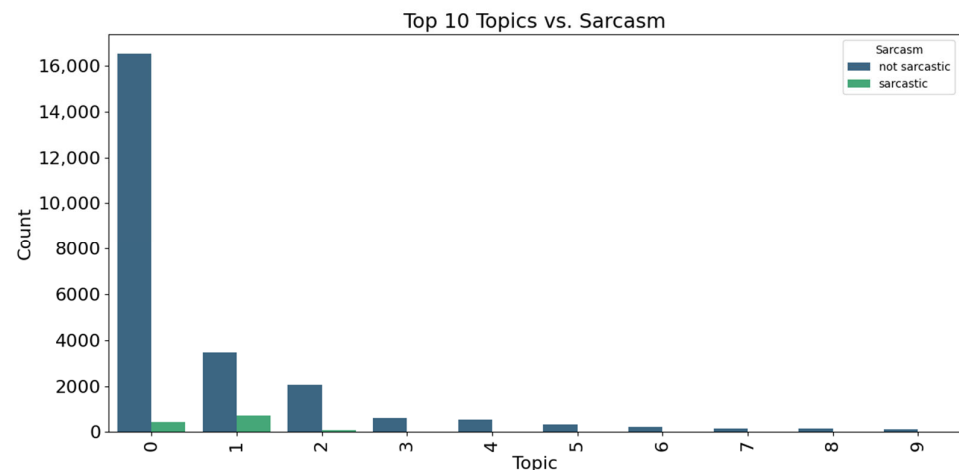


Figure 6. Distribution of sarcasm in the top 10 topics.

We can see that Topic 1, with the topic description “Requesting help creating a ChatGPT account in Egypt”, has the highest proportion of sarcastic tweets. Next is Topic 0, with the second-highest proportion of sarcastic tweets.

Table 4 below shows an example of sarcastic tweets in Topics 0 and 1. As Topic 1’s description indicates, the majority of sarcastic tweets are in the Egyptian dialect. This might be due to the MARBERT model being trained on a large portion of Egyptian tweets.

Table 4. Example of sarcastic tweets in Topics 0 and 1.

Topic 0	Topic 1
<p>واضح ان الحل الوحيد ان شات جي بي تي ليستبدلكش هو انك تشتغل رقاصه</p> <p>“It seems that the only way not be replaced by ChatGPT is to work as a belly dancer.”</p>	<p>اذكر واحد يقول شات جي بي تي بيستبدل الاخصائي النفسي دخلت سالتة ما جاوبني يقول منب مخول سلامات يا كوكو وانا مين يعالجني يا ذكاء اصطناعي</p> <p>“I remember someone saying that ChatGPT will soon replace psychiatrists. I asked ChatGPT and it didn’t reply saying ‘I am not authorized to answer’. OK AI, who will treat me then?”</p>
<p>سالت chatgpt عن المغرد كاسكو واجاب مغرد هطف</p> <p>“I asked ChatGPT about the tweeter, Casco, and it replied: a dumb tweeter”</p>	<p>وظائف ايه بس يا جماعه اللي يقضي عليها chatgpt و الراجل اللي بيقعد في الاسانسير يدوس علي الزرار لسه محافظ علي وظيفته في مصر لحد دلوقتي</p> <p>“What do you mean by ChatGPT eliminating jobs, when the elevator man in Egypt is still keeping his job.”</p>
<p>وضع الشعب مع chatgpt مثل وضع هالشايب مع توم القط</p> <p>“People and ChatGPT are like the situation of this old man with Tom the cat.”</p>	<p>بجد chatgpt دا شويه وهخليه يقوم يعمل اكل</p> <p>“Seriously, I will ask ChatGPT soon to prepare my meals.”</p>

5. Conclusions and Future Work

The aim of this study was to look at the primary topics discussed by early Arab users in Arabic tweets related to ChatGPT, the sentiments expressed in these tweets, and whether they convey a serious or sarcastic tone. To achieve these aims, a variety of data-analysis and topic-modeling techniques were used. Arabic tweets discussing ChatGPT use were collected using Python libraries *snsrape* and *tweepy*, focusing on specific keywords. The collected data underwent various cleaning and preprocessing steps, including removing punctuation, links, special characters, diacritics, emojis, and new lines. Duplicate and irrelevant tweets were also eliminated, resulting in a final dataset of 34,760 tweets.

The first objective of this study was to explore the primary topics discussed by early Arab users in Arabic tweets related to ChatGPT. Various topic-modeling algorithms, including LDA, NMF, and BERTopic, were tested. BERTopic was chosen for its superior performance on Arabic documents. This objective of the study was achieved, and the findings reveal a robust interest within the Arabic-speaking community regarding this technology, as evidenced by the prevalence of tweets discussing it. This popularity reflects ChatGPT's multidisciplinary influence and drives extensive social media discussions. The identified topics shed light on the dominant interest in ChatGPT technology within the Arabic-speaking community, as well as discussions around controversies, regional relevance, fake content and scams, and sector-specific dialogues. These findings seem consistent with other studies, e.g., [16,17,21]. However, alongside the enthusiasm, concerns, and controversies surrounding ethical risks and negative implications of ChatGPT emergence, it highlights apprehension surrounding advanced AI technology, particularly in language generation. Region-specific discussions in countries like Saudi Arabia, Libya, Egypt, and Morocco underscore the regional adoption and relevance of AI applications and ChatGPT technology in diverse contexts. The presence of topics related to fake ChatGPT apps and websites raises concerns about fraud and potential AI-generated content scams. Moreover, sector-specific dialogues exploring the impacts of ChatGPT across education, government, marketing, and employment demonstrate user curiosity and interest in understanding the potential benefits and challenges of integrating these technologies. According to [52], cultural and socioeconomic factors specific to each country can influence the adoption and discussions surrounding AI technology. Factors such as educational systems, technological infrastructure, government initiatives, and economic conditions may contribute to the varying levels of interest and engagement with ChatGPT technology in different regions.

The second objective of the present study involved exploring the sentiments expressed by early Arab users in their ChatGPT-related tweets. This objective of the study was achieved by employing a sentiment-analysis technique that was carried out with an unsupervised approach using the *ArabiTools* sentiment analysis. The analysis of 34,760 Arabic tweets classified into 50 topics revealed 32,267 neutral tweets (92.8%), 1348 positive tweets (3.9%), and 1145 negative tweets (3.3%). This indicated that the sentiment distribution was predominantly neutral, which is in line with other studies, e.g., [20,46]. The sentiment distribution within the first 10 topics showed that the majority of sentiments expressed across the topics were neutral, suggesting an inclination towards objectivity and factuality rather than being heavily influenced by subjective emotions. The neutral sentiment distribution indicates a focus on sharing information and exchanging knowledge rather than engaging in emotionally charged or opinionated debates. This suggests that early Arab users perceive ChatGPT as a tool for obtaining and disseminating information rather than a platform for emotional expression. Moreover, the prevalence of neutral sentiment suggests that objectivity and factuality are valued by early Arab users in their discussions related to ChatGPT. This implies a preference for evidence-based reasoning and logical arguments, which can contribute to more constructive and meaningful conversations. Cultural factors and communication norms within the Arab region can influence the expression of sentiments. Arab culture often emphasizes politeness, respect, and maintaining harmony in

discussions [53]. This cultural context may encourage users to adopt a neutral tone to avoid engaging in emotionally charged or opinionated debates.

The research also looked at sarcasm expressed by early Arab users in their ChatGPT-related tweets. This objective was achieved by analyzing the 34,760 Arabic tweets, which revealed that 1391 tweets (4%) are considered sarcastic. These sarcastic tweets were distributed across various topics, with Topic 1, focused on “Requesting help creating a ChatGPT account in Egypt,” having the highest proportion of sarcastic tweets. Topic 0, i.e., tweets related to Using ChatGPT, also had a significant number of sarcastic tweets. The prevalence of sarcasm in Topic 1 was primarily observed in the Egyptian dialect, while Topic 0 predominantly featured sarcastic tweets in the Saudi dialect. The low percentage of sarcastic tweets suggests that sarcasm may not be a dominant or frequently employed communication style in ChatGPT-related discussions among early Arab users. This could indicate that users prefer other forms of expression or that sarcasm may not be as prevalent or well-received in this particular context. It is important to note that the low percentage does not necessarily imply a complete absence of sarcasm but rather a lower frequency compared to other types of communication styles. Further research and analysis could provide additional insights into the factors influencing the use of sarcasm and its perceived effectiveness in ChatGPT-related conversations among Arab users.

The findings have several implications for AI developers, policymakers, and the Arab user community. The implications of the findings for AI developers are significant, as they highlight the robust interest and extensive social media discussions surrounding ChatGPT within the Arabic-speaking community. This indicates a strong demand for this technology and emphasizes the need for developers to address the concerns and controversies surrounding its ethical risks and negative implications.

AI developers should be aware of the apprehension surrounding advanced AI technology, particularly in the context of language generation. Our findings underscore the presence of concerns and controversies raised by users regarding the emergence of ChatGPT. This awareness should prompt developers to proactively address these concerns by prioritizing transparency, accountability, and ethical considerations in the development and deployment of language models. By demonstrating a commitment to responsible AI, developers can help alleviate user apprehension and build trust in the technology.

Additionally, our research findings emphasize the importance of educating Arab users about the ethical considerations, risks, and potential scams associated with AI chatbots. Developers should take proactive measures to provide clear and accessible information to users, empowering them to make informed decisions when engaging with AI-powered systems. This can include providing guidelines on responsible use, highlighting potential risks and limitations, and offering channels for users to report fraudulent or harmful activities.

Furthermore, the findings highlight the need for developers to mitigate potential negative impacts associated with ChatGPT. This can be achieved by implementing robust safeguards against biases, misinformation, and malicious use of language-generation technologies. Developers should invest in ongoing research and development to enhance the fairness, accuracy, and safety of AI models while actively addressing feedback and concerns raised by the community.

For policymakers, the findings provide insights into the regional adoption and relevance of AI applications and ChatGPT technology in diverse contexts, particularly in countries like Saudi Arabia, Libya, Egypt, and Morocco, where region-specific discussions were identified. These insights can inform policymakers’ understanding of the local dynamics and considerations related to AI technologies.

Firstly, this study highlights the regional adoption and relevance of AI applications and ChatGPT technology in different countries. Policymakers can use this information to assess the level of interest, engagement, and potential benefits of integrating AI technologies in their respective regions. Understanding the regional context is essential for policymakers to tailor policies, initiatives, and investments that support the responsible development, deployment, and utilization of AI technologies.

This study also raises concerns about fraud and potential AI-generated content scams within the Arabic-speaking community. Policymakers can leverage these findings to recognize the need for developing and implementing ethical guidelines and regulations that ensure responsible and accountable use of AI technologies. Such guidelines can address issues like data privacy, algorithmic transparency, and the prevention of malicious activities. By establishing clear guidelines and regulations, policymakers can promote the ethical and secure deployment of AI technologies, protecting users from potential harm and ensuring the integrity of AI-driven systems.

Furthermore, policymakers can consider this study's findings as an opportunity to engage in public discourse and consultation regarding the ethical, social, and legal aspects of AI technologies. By involving various stakeholders, including experts, industry representatives, and civil society organizations, policymakers can gather diverse perspectives and insights to inform the development of comprehensive policies and regulations. This inclusive approach can help address concerns, build public trust, and ensure that the adoption of AI technologies aligns with societal values and aspirations.

As for the Arab user community, this study sheds light on the preferences and characteristics of ChatGPT-related discussions among early Arab users. The findings suggest that there is a preference for objectivity and factuality in these discussions, as evidenced by the predominantly neutral sentiment distribution among the users.

This indicates that Arab users are more inclined towards sharing information and exchanging knowledge rather than engaging in emotionally charged or opinionated debates. The focus on objectivity suggests a desire for reliable and accurate information, fostering a culture of fact-checking and evidence-based reasoning. This preference for objectivity can contribute to more constructive and meaningful conversations where participants prioritize logical arguments and rational discourse.

The emphasis on evidence-based reasoning and logical arguments is a positive aspect of ChatGPT-related discussions within the Arab user community. It promotes critical thinking, encourages the evaluation of different viewpoints, and contributes to the cultivation of intellectual discourse. By valuing evidence and logical reasoning, users can engage in more productive discussions, leading to a deeper understanding of topics and fostering a culture of intellectual curiosity.

The study's findings also indicate that Arab users recognize the potential of ChatGPT as a tool for sharing information and knowledge. By leveraging ChatGPT and similar language-generation technologies, users can access a vast array of information, inquire about various topics, and engage in learning-oriented conversations. This highlights the value placed on knowledge exchange and the potential for ChatGPT to facilitate access to information and promote collaborative learning within the Arab user community.

Overall, the preference for objectivity, factuality, evidence-based reasoning, and logical arguments in ChatGPT-related discussions among Arab users suggests a focus on meaningful and constructive conversations. This inclination towards sharing information and exchanging knowledge fosters an environment that values intellectual discourse and promotes critical thinking. By recognizing these characteristics and preferences, the Arab user community can further harness the potential of ChatGPT for informed discussions, collaborative learning, and the dissemination of reliable information.

While this study aims to provide valuable insights into the topics, sentiments, and sarcasm expressed in ChatGPT-related tweets among early Arab users, it is important to acknowledge its limitations. One limitation is that the findings may be specific to the time period and context in which the data were collected. The rapidly evolving nature of online discourse and the continuous development of AI technologies could impact the generalizability of the conclusions over time. To address this, conducting longitudinal studies that span an extended period would help capture the temporal changes in sentiment expression and sarcasm over time.

Another limitation is the study's primary focus on ChatGPT-related tweets, potentially overlooking other platforms or modes of communication that may exhibit different patterns

of sentiment expression and sarcasm. To overcome this limitation, further studies are needed to explore the full range of sentiment and sarcasm expression in other digital spaces or offline interactions.

Future research involves conducting longitudinal studies that track sentiment expression and sarcasm over an extended period, which can provide a more comprehensive understanding of how these dynamics evolve over time. In addition, investigating sentiment and sarcasm expression beyond ChatGPT-related tweets, such as other social media platforms, online forums, or even offline interactions, can offer a broader perspective on user behavior and communication patterns. Future work may incorporate other modes of communication, such as audio or visual data, to capture additional layers of sentiment and sarcasm expression, as these can provide valuable insights into user interactions.

Author Contributions: Conceptualization, H.S.A.-K.; methodology, H.S.A.-K. and H.A.; software, S.A.-K. and F.A.; validation, S.A.-K. and F.A.; formal analysis, S.A.-K., F.A. and H.A.; investigation, S.A.-K. and F.A.; data curation, S.A.-K.; writing—original draft preparation, S.A.-K., F.A., H.A. and H.S.A.-K.; writing—review and editing, S.A.-K., F.A., H.A. and H.S.A.-K.; visualization, S.A.-K. and F.A.; supervision, H.S.A.-K.; project administration, H.S.A.-K.; funding acquisition, H.S.A.-K. All authors have read and agreed to the published version of the manuscript.

Funding: Researchers Supporting Project number (RSP2023R276), King Saud University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy and proprietary restrictions related to the use of Twitter data.

Acknowledgments: We acknowledge the use of ChatGPT, an AI chatbot developed by OpenAI, for generating some of the summaries in this article. ChatGPT was used to supplement our own writing and analysis, and not to replace them. We verified the accuracy and relevance of the AI-generated text before incorporating it into our manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Aljanabi, M. ChatGPT: Future Directions and Open Possibilities. *Mesopotamian J. Cybersecur.* **2023**, *2023*, 16–17. [[CrossRef](#)]
- Gill, S.S.; Kaur, R. ChatGPT: Vision and Challenges. *Internet Things Cyber-Phys. Syst.* **2023**, *3*, 262–271. [[CrossRef](#)]
- Biswas, S.S. Role of Chat Gpt in Public Health. *Ann. Biomed. Eng.* **2023**, *51*, 868–869. [[CrossRef](#)] [[PubMed](#)]
- Javaid, M.; Haleem, A.; Singh, R.P. ChatGPT for Healthcare Services: An Emerging Stage for an Innovative Perspective. *BenchCouncil Trans. Benchmarks Stand. Eval.* **2023**, *3*, 100105. [[CrossRef](#)]
- George, A.S.; George, A.H. A Review of ChatGPT AI's Impact on Several Business Sectors. *Partn. Univers. Int. Innov. J.* **2023**, *1*, 9–23.
- Li, L.; Ma, Z.; Fan, L.; Lee, S.; Yu, H.; Hemphill, L. ChatGPT in Education: A Discourse Analysis of Worries and Concerns on Social Media. *arXiv* **2023**, arXiv:2305.02201. [[CrossRef](#)]
- Lo, C.K. What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. *Educ. Sci.* **2023**, *13*, 410. [[CrossRef](#)]
- Rathore, B. Future of AI & Generation Alpha: ChatGPT beyond Boundaries. *Eduzone Int. Peer Rev./Ref. Multidiscip. J.* **2023**, *12*, 63–68.
- Kaddour, J.; Harris, J.; Mozes, M.; Bradley, H.; Raileanu, R.; McHardy, R. Challenges and Applications of Large Language Models. *arXiv* **2023**, arXiv:2307.10169.
- Zhuo, T.Y.; Huang, Y.; Chen, C.; Xing, Z. Red Teaming ChatGPT via Jailbreaking: Bias, Robustness, Reliability and Toxicity. *arXiv* **2023**, arXiv:2301.12867. [[CrossRef](#)]
- Ray, P.P. ChatGPT: A Comprehensive Review on Background, Applications, Key Challenges, Bias, Ethics, Limitations and Future Scope. *Internet Things Cyber-Phys. Syst.* **2023**, *3*, 121–154. [[CrossRef](#)]
- Alawida, M.; Mejri, S.; Mehmood, A.; Chikhaoui, B.; Isaac Abiodun, O. A Comprehensive Study of ChatGPT: Advancements, Limitations, and Ethical Considerations in Natural Language Processing and Cybersecurity. *Information* **2023**, *14*, 462. [[CrossRef](#)]
- Haleem, A.; Javaid, M.; Singh, R.P. An Era of ChatGPT as a Significant Futuristic Support Tool: A Study on Features, Abilities, and Challenges. *BenchCouncil Trans. Benchmarks Stand. Eval.* **2022**, *2*, 100089. [[CrossRef](#)]

14. Singh, D. ChatGPT: A New Approach to Revolutionise Organisations. *Ugc Approv. Res. J. India/UGC New. Added J./IJNMS* **2023**, *10*, 57–63.
15. Erfina, A.; Rifki Nurul, M. Implementation of Naive Bayes Classification Algorithm for Twitter User Sentiment Analysis on ChatGPT Using Python Programming Language. *Data Metadata* **2023**, *2*, 45. [\[CrossRef\]](#)
16. Korkmaz, A.; Aktürk, C.; Talan, T. Analyzing the User's Sentiments of ChatGPT Using Twitter Data. *Iraqi J. Comput. Sci. Math.* **2023**, *4*, 202–214. [\[CrossRef\]](#)
17. Haque, M.U.; Dharmadasa, I.; Sworna, Z.T.; Rajapakse, R.N.; Ahmad, H. "I Think This Is the Most Disruptive Technology": Exploring Sentiments of ChatGPT Early Adopters Using Twitter Data. *arXiv* **2022**, arXiv:2212.05856.
18. Tounsi, A.; Elkefi, S.; Bhar, S.L. Exploring the Reactions of Early Users of ChatGPT to the Tool Using Twitter Data: Sentiment and Topic Analyses. In Proceedings of the 2023 IEEE International Conference on Advanced Systems and Emergent Technologies (IC_ASET), Hammamet, Tunisia, 29 April 2023; pp. 1–6.
19. Birjali, M.; Kasri, M.; Beni-Hssane, A. A Comprehensive Survey on Sentiment Analysis: Approaches, Challenges and Trends. *Knowl.-Based Syst.* **2021**, *226*, 107134. [\[CrossRef\]](#)
20. Koonchanok, R.; Pan, Y.; Jang, H. Tracking Public Attitudes toward ChatGPT on Twitter Using Sentiment Analysis and Topic Modeling. *arXiv* **2023**, arXiv:2306.12951.
21. Leiter, C.; Zhang, R.; Chen, Y.; Belouadi, J.; Larionov, D.; Fresen, V.; Eger, S. ChatGPT: A Meta-Analysis after 2.5 Months. *arXiv* **2023**, arXiv:2302.13795. [\[CrossRef\]](#)
22. Zhou, J.; Müller, H.; Holzinger, A.; Chen, F. Ethical ChatGPT: Concerns, Challenges, and Commandments. *arXiv* **2023**, arXiv:2305.10646. [\[CrossRef\]](#)
23. Roumeliotis, K.I.; Tselikas, N.D. ChatGPT and Open-AI Models: A Preliminary Review. *Future Internet* **2023**, *15*, 192. [\[CrossRef\]](#)
24. Hassani, H.; Silva, E.S. The Role of ChatGPT in Data Science: How AI-Assisted Conversational Interfaces Are Revolutionizing the Field. *Big Data Cogn. Comput.* **2023**, *7*, 62. [\[CrossRef\]](#)
25. Liebreuz, M.; Schleifer, R.; Buadze, A.; Bhugra, D.; Smith, A. Generating Scholarly Content with ChatGPT: Ethical Challenges for Medical Publishing. *Lancet Digit. Health* **2023**, *5*, e105–e106. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Grassini, S. Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. *Educ. Sci.* **2023**, *13*, 692. [\[CrossRef\]](#)
27. Rahman, M.M.; Terano, H.J.; Rahman, M.N.; Salamzadeh, A.; Rahaman, M.S. ChatGPT and Academic Research: A Review and Recommendations Based on Practical Examples. *J. Educ. Manag. Dev. Stud.* **2023**, *3*, 1–12. [\[CrossRef\]](#)
28. Wen, J.; Wang, W. The Future of ChatGPT in Academic Research and Publishing: A Commentary for Clinical and Translational Medicine. *Clin. Transl. Med.* **2023**, *13*, e1207. [\[CrossRef\]](#)
29. Biswas, S. Prospective Role of Chat GPT in the Military: According to ChatGPT. *Qeios* **2023**. [\[CrossRef\]](#)
30. Carvalho, I.; Ivanov, S. ChatGPT for Tourism: Applications, Benefits and Risks. *Tour. Rev.* **2023**. *ahead-of-print*. [\[CrossRef\]](#)
31. Gursoy, D.; Li, Y.; Song, H. ChatGPT and the Hospitality and Tourism Industry: An Overview of Current Trends and Future Research Directions. *J. Hosp. Mark. Manag.* **2023**, *32*, 579–592. [\[CrossRef\]](#)
32. Tlili, A.; Shehata, B.; Adarkwah, M.A.; Bozkurt, A.; Hickey, D.T.; Huang, R.; Agyemang, B. What If the Devil Is My Guardian Angel: ChatGPT as a Case Study of Using Chatbots in Education. *Smart Learn. Environ.* **2023**, *10*, 15. [\[CrossRef\]](#)
33. Zhou, J.; Zhang, Y.; Luo, Q.; Parker, A.G.; De Choudhury, M. Synthetic Lies: Understanding AI-Generated Misinformation and Evaluating Algorithmic and Human Solutions. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 19 April 2023; pp. 1–20.
34. Santos, F.C.C. Artificial Intelligence in Automated Detection of Disinformation: A Thematic Analysis. *J. Media* **2023**, *4*, 679–687. [\[CrossRef\]](#)
35. Jones, V.A. Artificial Intelligence Enabled Deepfake Technology: The Emergence of a New Threat. Master's Thesis, Utica College, New York, NY, USA, 2020.
36. Sebastian, G. Do ChatGPT and Other AI Chatbots Pose a Cybersecurity Risk?: An Exploratory Study. *Int. J. Secur. Priv. Pervasive Comput.* **2023**, *15*, 1–11. [\[CrossRef\]](#)
37. Derner, E.; Batistič, K. Beyond the Safeguards: Exploring the Security Risks of ChatGPT. *arXiv* **2023**, arXiv:2305.08005. [\[CrossRef\]](#)
38. Van Dis, E.A.; Bollen, J.; Zuidema, W.; van Rooij, R.; Bockting, C.L. ChatGPT: Five Priorities for Research. *Nature* **2023**, *614*, 224–226. [\[CrossRef\]](#)
39. Abdullah, M.; Madain, A.; Jararweh, Y. ChatGPT: Fundamentals, Applications and Social Impacts. In Proceedings of the 2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS), Milan, Italy, 29 November 2022; pp. 1–8.
40. Skjuve, M.; Følstad, A.; Brandtzaeg, P.B. The User Experience of ChatGPT: Findings from a Questionnaire Study of Early Users. In Proceedings of the 5th International Conference on Conversational User Interfaces, Eindhoven, The Netherlands, 19 July 2023; pp. 1–10.
41. Xu, R.; Feng, Y.; Chen, H. ChatGPT vs. Google: A Comparative Study of Search Performance and User Experience. *arXiv* **2023**, arXiv:2307.01135. [\[CrossRef\]](#)
42. Salah, M.; Alhalbusi, H.; Ismail, M.M.; Abdelfattah, F. Chatting with ChatGPT: Decoding the Mind of Chatbot Users and Unveiling the Intricate Connections between User Perception, Trust and Stereotype Perception on Self-Esteem and Psychological Well-Being. *Curr. Psychol.* **2023**. [\[CrossRef\]](#)

43. Beltrami, E.J.; Grant-Kels, J.M. Consulting ChatGPT: Ethical Dilemmas in Language Model Artificial Intelligence. *J. Am. Acad. Dermatol.* **2023**, S019096222300364X. [[CrossRef](#)]
44. Sharma, S.; Aggarwal, R.; Kumar, M. Mining Twitter for Insights into ChatGPT Sentiment: A Machine Learning Approach. In Proceedings of the 2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), Ballar, India, 29 April 2023; pp. 1–6.
45. Taecharungroj, V. “What Can ChatGPT Do?” Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data Cogn. Comput.* **2023**, *7*, 35. [[CrossRef](#)]
46. Kherwa, P.; Bansal, P. Topic Modeling: A Comprehensive Review. *ICST Trans. Scalable Inf. Syst.* **2018**, *7*, 159623. [[CrossRef](#)]
47. Sarkar, T.; Rajadhyaksha, N. TLA: Twitter Linguistic Analysis. *arXiv* **2021**, arXiv:2107.09710. [[CrossRef](#)]
48. Chaudhary, J.; Niveditha, S. Twitter Sentiment Analysis Using Tweepy. *Int. Res. J. Eng. Tech.* **2021**, *8*, 4512–4516.
49. Grootendorst, M. BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure. *arXiv* **2022**, arXiv:2203.05794. [[CrossRef](#)]
50. Abdelrazek, A.; Medhat, W.; Gawish, E.; Hassan, A. Topic Modeling on Arabic Language Dataset: Comparative Study. In *International Conference on Model and Data Engineering*; Springer Nature: Cham, Switzerland, 2022; pp. 61–71. [[CrossRef](#)]
51. Rahma, A.; Azab, S.S.; Mohammed, A. A Comprehensive Review on Arabic Sarcasm Detection: Approaches, Challenges and Future Trends. *IEEE Access* **2023**, *11*, 18261–18280. [[CrossRef](#)]
52. Ajaaj, M.A.-Q. Politeness Strategies in Arabic Culture with Reference to Eulogy. *EFL J.* **2016**, *1*, 161–173. [[CrossRef](#)]
53. Vu, H.T.; Lim, J. Effects of Country and Individual Factors on Public Acceptance of Artificial Intelligence and Robotics Technologies: A Multilevel SEM Analysis of 28-Country Survey Data. *Behav. Inf. Technol.* **2022**, *41*, 1515–1528. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.