An In-depth Analysis of Current Trends in ChatGPT Academia, Scholars' Sentiments, and Network Characteristics in Co-authorship of Published Papers

Lingran Hu[1,a], Qinxin Xu[2,b], Yingchen Yao[3,c,*], Tongwei Zhang[4,d]

[1]*Wakeforest University, Winston-Salem, NC 27109, US*
[2]*Crofton House School, Vancouver, V6N3E1, Canada*
[3]*SPSS, The University of Edinburgh, Edinburgh, EH165AY, UK*
[4]*Ameican Heritage, Florida, 33026, US*
a. lingranhu@outlook.com, b. lisax0228@gmail.com,
c. 937335038mist@gmail.com, d. pl250951@ahschool.com
*corresponding author

Abstract: This paper provides a comprehensive analysis of the current trends in ChatGPT academia, scholars' sentiments, and network characteristics in the co-authorship of published papers. With the rapid development and widespread adoption of ChatGPT, a powerful AI language model developed by OpenAI, there is a growing need to understand its impact on academia and society at large. This work aims to address three main questions: (1) What are the hot topics researchers are exploring when it comes to ChatGPT? (2) How do scholars feel about this new technology? (3) What can we learn from looking at who collaborates on academic papers about ChatGPT? Utilizing techniques like MDCOR and the R language, this work meticulously analyzes academic results, public sentiments, and co-authorship networks. The findings highlight key themes and sentiments in the academic community, revealing a generally positive attitude towards ChatGPT, but with concerns about potential risks and limitations. Additionally, the analysis of co-authorship networks underscores the need to promote diverse collaborations to avoid education monopolies. In conclusion, this study offers valuable insights for responsible AI integration and academic progress, shedding light on the prevailing trends, sentiments, and network characteristics in ChatGPT academia.

Keywords: ChatGPT, Academic Field Analysis, Sentiment Analysis, Co-authorship Network, Artificial Intelligence Applications

1. Introduction

Founded in 2015, OpenAI carries the goal to improve the safety of artificial intelligence. On November 30, 2022, they unveiled ChatGPT, a powerful AI language model. Imagine a chatbox where you can have interactive conversations with an AI - that's ChatGPT. It's now gathering feedback from users worldwide, helping developers fine-tune and enhance it. Impressively, ChatGPT garnered over a million users in its first five days and hit 100 million monthly active users just two months after launch, as reported by Forbes in 2023 [1]. ChatGPT has strengths, like quick and accurate responses while avoiding controversial topics. However, it also has its weaknesses. Some

worry that its knowledge base is limited and prone to biases. Misinformation is another concern, given its massive user base.

This paper will be oriented around three main questions

● **Trends in Academic Research:** What are the hot topics researchers are exploring when it comes to ChatGPT?

● **Scholar Sentiments:** How do scholars feel about this new technology?

● **Network Characteristics:** What can we learn from looking at who collaborates on academic papers about ChatGPT?

This paper digs into public feelings about ChatGPT, aiming to figure out if the overall sentiment is positive or negative. Using techniques like MDCOR and the R language, it meticulously analyzes academic works. In the next section, the paper reviews existing literature, pointing out gaps and revealing its own chosen paths. Then, it lays out the data and methods used for analysis in Section 3. Progressing to Section 4, it visually presents the findings, highlighting key themes and sentiments. Section 5 takes on implications, limitations, and ethical concerns, including an exploration of the concept of education monopoly, as discussed by Nielsen in 2022 [2]. By tapping into scholarly insights, the paper aims to cast light on where ChatGPT is headed and how it affects society at large.

2. Literature Review

2.1. Summary of Papers

This literature review compiles research papers from Google Scholar to provide a deeper understanding of ChatGPT's impact. These papers investigate how ChatGPT affects academia, including scholars, teachers, students, and the broader societal future. Current research on ChatGPT largely falls into three categories:

● **General Overview:** Some scholars offer a comprehensive understanding of ChatGPT by discussing its capabilities, limitations, and potential for the future. These studies provide an overall picture of the technology's implications without focusing on specific areas.

● **Public Opinion Analysis:** Other studies delve into the public's perceptions of ChatGPT. Researchers collect data from social media platforms to gauge whether the public views this technology positively or negatively. This approach helps individuals understand how society at large receives ChatGPT.

● **Domain-Specific Applications:** Another group of scholars examines how ChatGPT is applied in specific fields. These studies highlight how ChatGPT benefits professionals in various sectors. By looking at its impact in these domains, researchers offer insights into how ChatGPT enhances specific areas of work.

By categorizing research into these three areas, scholars can develop a well-rounded understanding of how ChatGPT influences academia and society.

2.2. Specific Areas

One of the initial papers under consideration is titled "Summary of ChatGPT/GPT-4 Research and Perspective Towards the Future of Large Language Models" [3]. This work draws extensively from prior research and employs data modeling to visually represent the distribution of papers across various fields. It not only assesses ChatGPT's effectiveness, spotlighting its strengths and limitations, but also projects its potential applications in the future.

Another contribution to this discourse comes from Ray, who provides a comprehensive overview of ChatGPT's capabilities, challenges, and potential. The discussion includes an exploration of ChatGPT's utility in the scientific realm, encompassing tasks such as "data processing, hypothesis generation, and collaboration." In contrast to the paper, Ray introduces discussions about various

technological models, thus providing a more encompassing perspective. Both these references serve as valuable resources for scholars seeking to delve into ChatGPT's multifaceted capabilities [4].

Following a similar structure, this paper begins by introducing ChatGPT, outlining the research objectives, describing data collection methods, and culminating in an exploration of ChatGPT's practical applications.

On the other hand, Qi et al. analyzed public perceptions concerning the increasing usage of ChatGPT, specifically within social media platforms such as subreddits. Their investigation centers around two pivotal questions:

1) What are the specific topics being discussed in relation to ChatGPT?
2) What are the dominant sentiments associated with these topics?

Through the utilization of topic modeling, cross-subreddit analysis, and sentiment analysis, their study highlights a generally positive public perception of ChatGPT. In this paper, R-studio is adopted for sentiment analysis, taking a more nuanced approach by delving beyond the dichotomy of positive and negative sentiments. Specific emotions such as anger, anticipation, and excitement will be explored, allowing the research group to understand the unique perspectives of scholars and gain insights into ChatGPT's potential trajectory [5].

Lastly, the third category of research investigates ChatGPT's role within specific domains. Biswas, for example, underscores ChatGPT's potential in "advancing our understanding of climate change and enhancing the accuracy of climate projections." The paper delineates tasks such as "model parameterization, data analysis and interpretation, scenario generation, and model evaluation" [6]. While acknowledging potential challenges like comprehending complex scientific concepts, contextual awareness gaps, and potential bias in training data, these studies effectively bridge the gap between technology and practical applications. This alignment with practical implications resonates with this paper's broader purpose [6].

2.3. Gaps

This paper brings fresh perspectives to the ongoing ChatGPT discourse by bridging gaps and offering a distinctive approach that sets it apart from earlier scholarly work.

To start, the paper leverages the MDCOR framework for analysis, providing an in-depth understanding of the topics covered, prevalent terms, and more. The data section will elaborate on this, offering insights into the numerical representation within the dataset.

A significant departure from prior studies is this paper's incorporation of sentiment analysis and a comprehensive overview. Building on the extended timeline since ChatGPT's inception, this approach integrates these dimensions. In addition to outlining scholars' areas of interest through Scopus, the paper delves into their perceptions of ChatGPT.

The primary focus remains on scholars as key stakeholders, meticulously dissecting their viewpoints. This approach offers benefits like accuracy and relevance, courtesy of scholars' profound expertise. Simultaneously, it introduces potential limitations in sample representativeness due to the specialized nature of this group.

In summary, researchers pursue three avenues: examining ChatGPT itself, gauging public opinions, and evaluating ChatGPT's role in specific fields. Scholars express concerns about ChatGPT's ethical performance and privacy implications. For example, Liu et al. emphasize ChatGPT's inconsistent moral stance, demonstrated by varied responses to the trolley problem, which could have negative societal ramifications. Furthermore, concerns about privacy emerge due to ChatGPT's utilization of public data. Nevertheless, scholars identify a prevailing positive sentiment among the public toward the current ChatGPT model, anticipating an optimistic path for its future development.

3. Method

In this study, three distinct methodologies were employed, utilizing MDCOR (Machine Driven Classification of Open-Ended Responses) and R-Studio as bases for NLP (Natural Language Processing), generating analytics of sentiments and social network structures.

3.1. Data Source

The primary data source for this study was the Scopus database. Scopus developers claim to index over 14,000 STM and social science titles from 4000 publishers [7]. This vast collection ensures representation across various fields, alleviating concerns regarding any lack of inclusivity or underrepresentation. The reason this study chose Scopus is due to its comprehensive coverage of academic literature and relevance to the research question in the academic field. Approach For the data analysis, Digital Object Identifiers (DOIs) and abstracts were chosen as the focus. A digital object identifier (DOI) is a persistent identifier or handle used to uniquely identify various objects, standardized by the International Organization for Standardization (ISO), enabling accurate and precise individual paper analysis [8]. Abstracts, on the other hand, provide succinct summaries of the main content, facilitating efficient data analysis while reducing the need to examine the entire paper. Focusing on DOIs and abstracts also served to mitigate potential copyright concerns, as substantial portions of the papers were not reproduced in our analysis.

3.2. Data Selection

This study selected academic papers that were relevant to ChatGPT, focusing on those published from January 1st, 2023, to the present. Due to the inability to differentiate between GPT-3.5 and GPT-4.0 in the available data, a time frame comparison was deemed unfeasible and not undertaken.

3.3. Data Cleaning

Among the extracted 897 papers from Scopus, a subset was found to lack abstracts. To ensure data integrity and avoid any ambiguity or interference, these papers were excluded from our analysis using spreadsheet software, resulting in a final dataset comprising 548 papers.

3.4. Data Analysis Method with MDCOR

To categorize and classify the dataset effectively, the MDCOR (Machine Driven Classification of Open-Ended Responses) method was employed. MDCOR is a machine learning-driven text classification approach that utilizes data analysis, data visualization, and Natural Language Processing (NLP) techniques to efficiently categorize open-ended questions [9]. MDCOR was chosen for its ability to quickly and accurately categorize open-ended responses, enabling a systematic examination of the research question's various aspects. The method's reliance on machine learning and data-driven analysis makes it suitable for addressing Research Question 1, which centers on identifying the diverse topics scholars are focusing on regarding ChatGPT.

3.4.1. MDCOR Execution

The data analysis execution involved several steps to process and analyze the dataset using the MDCOR (Machine Driven Classification of Open-Ended Responses) method.

Firstly, Figure 1 shows that text cleaning and mining procedures were conducted to prepare the dataset for analysis. Secondly, metrics (Figure 1) were executed to determine the optimal number of

codes, employing four convergence metrics to assess agreement. The convergence of the two metrics indicated that 13 codes constituted the most optimal solution for classifying the dataset.

The final step involved selecting the optimal number of codes for the machine-learning process. The value was set to 13 based on the metrics assessment results, which indicated that each code represented distinct content and was discernible from the others.

The MDCOR was executed with the selected thirteen-code solution. Figure 2 shows that the analysis generated an interactive descriptive tool that quantified word frequency distributions per code and indicated the top 30 most essential words for each code (Figure 2). Additionally, the tool displayed distances among codes in a quadrant plane to convey similarities or differences in underlying meanings.

The fully classified dataset contained all codes and texts, allowing researchers to understand the associations between textbooks and codes probabilistically. Measures of text contribution and group fit were provided to assess text-to-code strength and detect the most representative texts per code.
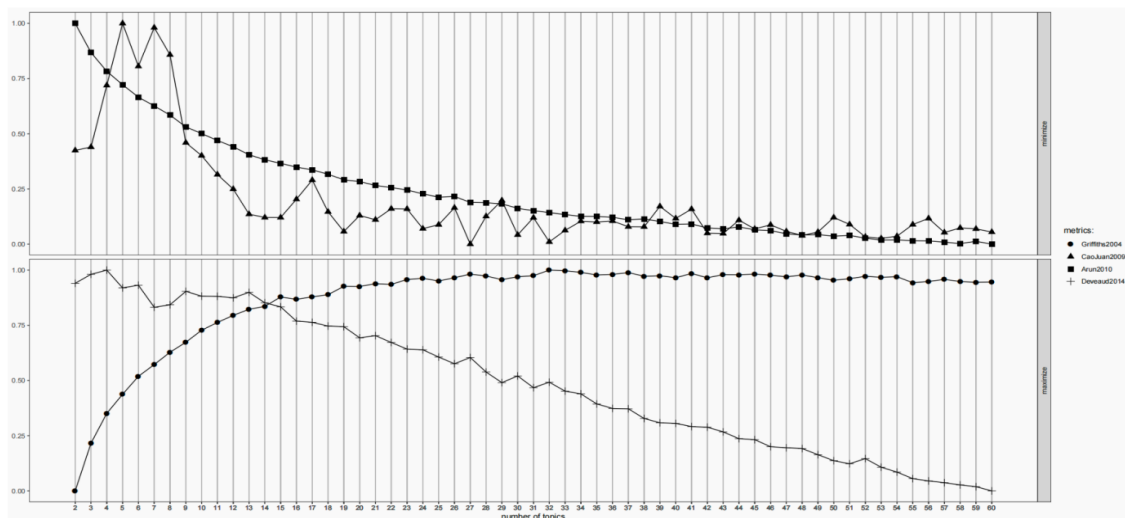

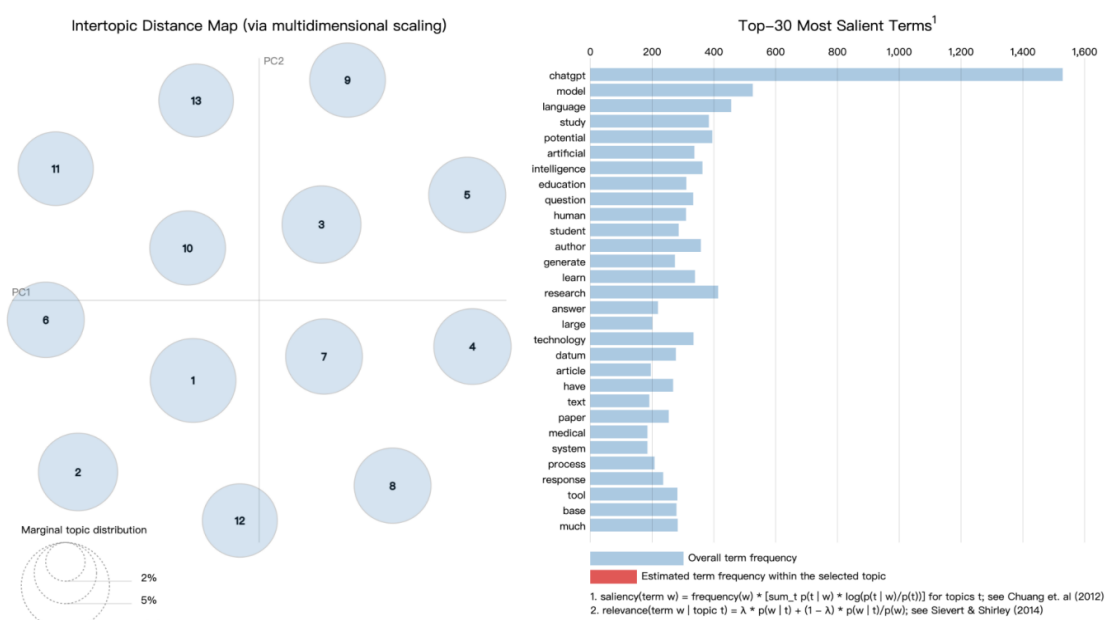Figure 1: Metrics output generated through MDCOR

Figure 2: Intertopic distance map generated through MDCOR

3.5. R-Studio for Sentimental Analysis

```
1  library(tm)
2  library("textstem")
3  library("ggplot2")
4  library("plotly")
5  library("reshape2")
6  library("syuzhet")
7  library("stringr")
8
9  #corpus a collection of documents
10 x1<-read.csv("/Users/chriszhang/Desktop/CIS/CCCCHATGPT CSV.csv")
11   x2<-x1[, c("DOI", "Abstract")]
12   names(x2)<-c("doc_id", "text")
13   x2$text <- str_remove_all(x2$text, "[^[\\da-zA-Z ]]")
14   docs <- Corpus(DataframeSource(x2),readerControl = list(language = "en"))
15
16
17 #Transformations
18 docs <- tm_map(docs, content_transformer(tolower))
19 docs <- tm_map(docs, removeNumbers)
20 docs <- tm_map(docs, removeWords, stopwords("english"))
21 # To remove stopwords who what when where why does can can't doesn't didn't and how could should
   didnt
22 newstopwords <-c("with","and", "for", "the", "to", " in ", "then", "he", "she", "than", "not",
   "that", "thats", "dont", "whats","you","for","are","but","have","your","this","was","like","just"
   ,"all","its","will","there","get","they","from","one","about","some","would","more","them","really"
   ,"had","been","also","because","did","has","which","were","other","very","into","her","their","even"
   ,"his","say","ive","youre","(",")","httpwwwfacebookcomwethingtonphotography","i've","it's",":","i'm"
   ,"i'll","---","don't","and", "i'd","won't","aren't","wasn't","we'll", "ill","cant", "can't",
   "yourownwords")
23 docs <- tm_map(docs, removeWords, newstopwords)
24 docs <- tm_map(docs, removePunctuation)
25 docs <- tm_map(docs, stripWhitespace)
26
27 #Normalization #Think about whether you should do this
28 #docs <- tm_map(docs, lemmatize_strings)
29
30 # SENTIMENTS FREQUENCY
31
32 sentiments_db <- get_nrc_sentiment(docs$content, cl = NULL, language = "english", lowercase = TRUE)
33 rownames(sentiments_db)<- names(docs$content)
34
35 sentimentScores = data.frame(colSums(sentiments_db))
36
37 names(sentimentScores) = "Score"
```

Figure 3: R-Studio work-in-progress for sentiment analysis

For the following sentiment analysis and SNA (social network analysis), R-Studio will be applied, serving as the basis for the NLP (natural language processor) analysis process. After analyzing the degree of sentiments in those papers shown in Figure 3, eventually compiled them into a congregated sentiment analysis. Different from the analysis of the current trends, R-Studio [10] is used as the basis for an NLP (natural language processor) to analyze the choice of words from the papers. In this text mining process, first, remove all the emotionless unbiased words that aren't in the lexicon, then locate the corresponding sentimental words in the papers. The referencing lexicon is the NRC Word-Emotion Association Lexicon, which contains eight basic emotions that are commonly applicable to NLP and sentiment analysis analyses [11].

3.6. R-Studio for Co-authorship Analysis

```
E(g)$weight <- 1 #must step
g.c <- simplify(g)
E(g.c)$weight

#centrality measures
cent<-data.frame(ID=V(g.c)$name, ev=evcent(g.c)$vector, deg=degree(g.c)/max(degree(g.c)), bet
=betweenness(g.c, normalized=F)/max(betweenness(g.c, normalized=F)), clo=closeness(g.c)/max
(closeness(g.c)))
head(cent)
cent$max_cent <- cent$ev + cent$deg + cent$bet + cent$clo

cent<- cent[order(cent$max_cent, decreasing=T),]
head(cent)

library(networkD3)
V(g.c)$label <- V(g.c)$name
V(g.c)$name<-1:length(V(g.c)) #(1:length(V(g)))-1
links<-as.data.frame(cbind(get.edgelist(g.c),as.numeric(E(g.c)$weight)))

links$V1<-as.numeric(as.character(links$V1))
links

links$V2<-as.numeric(as.character(links$V2))
str(links)
```

Figure 4: R-Studio work-in-progress for sociogram

To further explore the interconnection among the co-authors of the research papers, R-Studio is again the basis for the social network analyzing NLP shown in Figure 4, eventually generating a sociogram that highlights the most active individuals, known as the "Actors" [12]. The mathematical representation of the vertex implemented to analyze sociograms is G = (E, V) where E is the edges/links/connects/relations whereas V is the units/actors/vertices/nodes for which is being observed. It also depends on the type of relation analyzing, the definition of E will change correspondingly to different scenarios. After analyzing researchers' opinions on Chat GPT the next is the co-authorship network, which is how people are connected through published papers on the topic of Chat GPT. The sociography generated locates the most relevant actors to their field and also their centrality measures which are based on several different degrees. The Degree of centrality counts the connections among the co-authors, which is how many times one author publishes with someone else. Betweenness centrality is how often a given actor is between two other actors, functioning as a connection between people. Eigenvector centrality/Mathew effect, in simple terms, is advantage begets advantage; those having the most continue to have the most. In co-authorship analysis, an author will be relevant to the extent that his/her connections are relevant, if he/she is friends with an actor then he/she is also more likely to be active in the sociogram. Lastly, closeness centrality, computed as marginals of symmetric geodesic distance matrix; the distance between the co-author of one's co-author, the lower the value, the more active an author is in the sociogram. If a given actor has a value of 1 in all these centrality measures, then the actor would have an overall centrality of 4, making him/her the most influential/active among them all.

4. Results

4.1. Data Analysis with MDCOR Results 1

The MDCOR analysis identified key topics related to ChatGPT. The integrated Excel table (Table 1) provided a visual representation of the dominant keywords for each topic. Distinct themes were explored, including potential benefits and challenges, data analysis applications, ethical considerations, intelligent systems, conversational software, deep querying solutions, and societal applications shown in Table one. The analysis also covered evaluation, law, integration's impact, text generation quality, time efficiency, programming ability, authorship, and science and technology

implications. Overall, the MDCOR analysis offers valuable insights for the field of natural language processing and conversational AI.

Table 1: Excel table with keywords for each topic

Top-30 Most Relevant Terms for Topics/Number of Topics	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	chatgpt	chatgpt	potential	education	medical	datum	chatgpt	study	model	technology	generate	human	author	intelligence
2	model	question	include	student	patient	system	application	chatgpt	language	paper	text	have	article	artificial
3	language	answer	research	learn	health	base	society	analysis	large	information	chatbot	good	research	author
4	study	response	limitation	assessment	healthcare	user	engineer	chatbots	process	research	response	much	scientific	generative
5	potential	provide	ethical	tool	clinical	approach	openai	much	natural	design	prompt	like	academic	tool
6	artificial	performance	concern	high	make	machine	develop	result	lims	future	write	problem	review	f
7	intelligence	evaluate	challenge	teach	care	propose	train	relate	paper	provide	quality	code	publish	science
8	education	result	benefit	educational	improve	framework	public	social	train	limit	also	think	literature	technology
9	question	compare	issue	technology	risk	lean	also	examine	task	knowledge	high	work	write	advance
10	human	accuracy	need	challenge	decision	technique	licence	finding	case	service	base	program	have	many
11	student	test	address	write	medicine	intelligent	exclusive	level	capability	explore	create	current	work	field
12	author	report	explore	practice	professional	interaction	industry	find	common	development	good	understand	journal	become
13	generate	search	datum	teacher	practice	discuss	attention	effect	present	group	ability	however	researcher	communication
14	learn	information	field	support	provide	method	biomedical	conduct	generative	library	produce	ability	publication	implication
15	research	correct	however	digital	treatment	integrate	several	medium	demonstrate	purpose	across	recent	nature	change
16	answer	score	discuss	educator	evidence	development	launch	identify	open	impact	specific	task	authorship	community
17	large	google	application	university	personalize	enable	transformer	utilize	domain	value	form	year	springer	management
18	technology	rate	highlight	skill	support	develop	generation	significantly	automate	area	detect	help	first	material
19	datum	conclusion	offer	suggest	diagnosis	conversational	recently	increase	license	process	find	still	part	emerge
20	article	examination	impact	course	however	software	since	investigate	legal	implication	note	computer	policy	application
21	have	different	various	provide	base	build	release	evaluation	apply	finding	detection	solve	discussion	rapidly
22	text	perform	development	learner	source	much	november	focus	general	present	algorithm	life	editorial	world
23	paper	exam	ensure	require	potential	concept	particularly	perception	international	improve	output	need	consider	example
24	medical	give	opportunity	peer	surgery	virtual	power	reveal	standard	taylor	personal	robot	guideline	regard
25	system	much	bias	feedback	revolutionize	allow	late	effectiveness	creative	methodology	essay	show	toal	give
26	process	multiple	future	outcome	outcome	complex	capability	positive	copyright	integration	statement	although	therefore	scenario
27	response	objective	consideration	integrity	plan	solution	chat	self	specific	francis	time	require	reference	license
28	toal	f	associate	institution	dataset	platform	million	influence	attribution	involve	respond	reason	editor	perspective
29	base	term	range	create	enhance	query	make	specifically	transformer	focus	summary	possible	must	conclude
30	much	average	privacy	experience	physician	deep	effectively	analyze	summarizatior	understand	internet	especially	right	current

4.2. Data Analysis with Sentiment Results 2

After mining the text corpus, an intuitive graph will be generated by R-Studio, which is placed below. Due to the tendency to stay impartial in research papers, the two most displayed sentiments are "positive" and "trust", with the third most displayed in anticipation since it's a newly surfacing product. However, pessimistic sentiments still exist, such as "negative", "anger" and "fear" as shown in Figure 5. The shown sentiments act as a somewhat subtle complementary to the overwhelming sentiment of "positive", insinuating that the academic area as a whole exemplifies mixed feelings towards Chat GPT to a considerable degree.
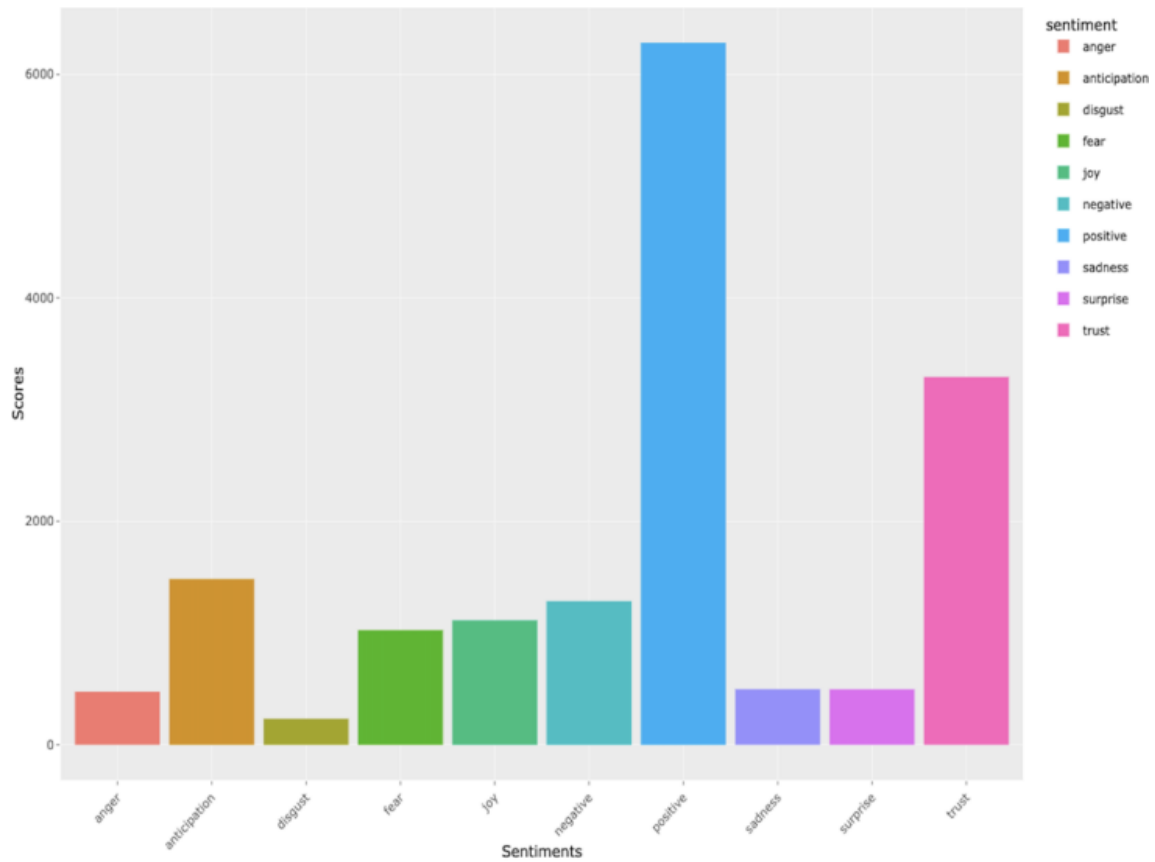
Figure 5: Sentiment Analysis generated through R-Studio

4.3. Data Analysis with SNA Results 3

Through data processing, the located author "Sascha Kraus" is shown to be the most active "Actor" with an overall centrality measure of approximately 2.989, as shown in Table 2. Sascha Kraus, as an active author in the fields of entrepreneurship and business, has stretched his influence to the area of artificial intelligence. Currently, hold the position of Full Professor of Management at the Free University of Bozen-Bolzano located in Italy, with another title as a Distinguished Visiting Professor in Entrepreneurship and Digital Transformation at the University of Johannesburg in South Africa [13].

Table 2: Top 6 most influential actors identified through R-Studio

	ID	ev	deg	bet	clo	max_cent
57207720523	57207720523	0.9863187	1	9.91E-01	0.01136364	2.988944
9633912200	9633912200	0.9859162	0.9736842	5.08E-01	0.01111111	2.479072963
41961835900	41961835900	0.9859162	0.9736842	5.08E-01	0.01111111	2.479072
57813546700	57813546700	1	0.9736842	2.51E-01	0.01098901	2.235655
35239818900	35239818900	1	0.9736842	2.51E-01	0.01098901	2.235655
57190980898	57190980898	0.9855361	0.9473684	4.293222-05	0.01086957	1.943817

5.   Discussion

5.1.   Discussion of Data

5.1.1. Topic Exploration

The identified research topics shed light on the multidimensional aspects of ChatGPT research in academia. Topics such as the impact of artificial intelligence, advancements in technology, and ethical considerations emerged as focal points of interest.

5.1.2. Challenges in Defining Optimal Topic Quantity

The metrics plot, though informative, did not yield a singular optimal number of topics. Nonetheless, a comparative analysis was undertaken to maximize code differences, leading to robust research insights.

5.1.3. Limitation of Research Papers

This study has focused on various aspects of ChatGPT academia, including current trends, scholars' sentiments, and network characteristics in the co-authorship of published papers. By analyzing these sentiments, we aim to provide a comprehensive representation of the academic field's overall attitude towards ChatGPT, which can be inferred to reflect the AI industry's perception of ChatGPT's distinctiveness.

However, it is important to acknowledge that the data used in this study primarily comes from scholars and academic publications. While the academic field holds significant influence and is often considered a leading voice, it may not fully represent the perspectives of the general public. This limitation arises from the lack of representation from the main public, whose opinions and perceptions may differ from those of scholars.

5.2.   Implications

It is crucial to recognize that the results and opinions conveyed through research papers can impact the public's viewpoint to an extent. Therefore, as more academic papers on ChatGPT emerge, it becomes increasingly important to capture the influences of the perspectives of the broader public experience. This will help provide a more comprehensive understanding of ChatGPT's reception and its implications for society.

5.2.1. Implications of knowing the trends of research topics in the academic field of ChatGPT

The findings from the present study indicate that ChatGPT demonstrates significant potential in 13 key domains: technology, education, medicine, and law…. therefore offering implications for various areas, including investment, regulation, and further research.

From an investment perspective, the promising prospects of ChatGPT extend beyond the immediate domain of language generation [14]. Related fields, such as electronic health records (EHRs) in the medical sector [15], ChatGPT's integration into front-end technology operations to offer banking services directly to customers [16] and the development of ChatGPT Adaptive Learning systems that are capable of personalized teaching methods based on individual student progress and performance [17] all present promising investment opportunities.

In terms of regulation, the increasing usage of AI and ChatGPT in healthcare communication [18] raises concerns regarding potential risks of misinterpretations that could negatively impact patient well-being. Policymakers should proactively develop guidelines and monitoring mechanisms to

ensure patient safety and minimize chances of miscommunication. Additionally, the integration of AI technologies in academic contexts raises questions about academic honesty and plagiarism [19]. Implementing regulations that address these issues will be essential to uphold academic integrity and ensure ethical AI deployment. Furthermore, policymakers could revisit the risk-based AI taxonomy outlined in the AI Act to verify that the regulatory framework remains up-to-date and aligns with the evolving landscape of AI technologies, including ChatGPT [20].

In terms of future research, it is imperative for researchers to continually explore new problems and propose creative solutions to advance academic innovation [21]. Given the demonstrated potential of ChatGPT in technology, education, and the medical field, directing research efforts toward these domains can lead to significant advancements that benefit society at large.

5.2.2. Implications of knowing scholars' sentiments towards ChatGPT

In regards to sentiment analysis, although the scholarly response to ChatGPT has predominantly been positive and trusting, it is essential to acknowledge and address the potential negative aspects associated with its use. Such apprehensions should prompt a more comprehensive approach to building relationships and enacting laws that mitigate possible technological risks. For instance, there are concerns about the possibility of cheating on online exams facilitated by ChatGPT's human-like text generation capabilities, which necessitates measures to ensure academic integrity and maintain the credibility of online assessments [22]. Moreover, the research highlights the potential challenges posed by ChatGPT in areas such as diminished critical thinking skills and difficulties in evaluating information generated by the system. These issues underscore the need for education and awareness to promote critical thinking and discernment among users when engaging with content produced by ChatGPT. Furthermore, these negative sentiments raise important considerations related to using ChatGPT for safety-related information. It is important to have cautious reliance on ChatGPT for critical matters where human expertise and verification are essential to ensure accuracy and reliability [23]. Building a robust framework for responsible AI usage will be vital in fostering trust and confidence among users and stakeholders.

5.2.3. Implications of knowing the network characteristics of published academic papers

The implications of understanding the network characteristics of published academic papers extend beyond the immediate scope of the present research. By comprehending and analyzing these characteristics, researchers can gain valuable insights that have the potential to prevent potential education monopolies. Such monopolies can have far-reaching consequences, including the establishment of centralized bureaucratic systems within academia [24]. These bureaucratic structures may impede the free flow of knowledge and ideas, leading to pedagogically stagnant environments that hinder innovation and progress in education.

While the current research on ChatGPT is relatively new, and there may not be a clear monopoly trend evident at this stage, the methodology employed in this study holds promise for addressing this concern in the future. By leveraging the knowledge of network characteristics in academic publishing, the research can aid in fostering a more open and competitive academic landscape, encouraging diverse perspectives and fresh ideas to flourish.

Identifying the specific active author as a professor at Italy's University of Johannesburg, not only provides insights into the individual's expertise and research contributions but also offers valuable information about the region and the university's current leadership and advancements in the field of ChatGPT.

This knowledge can have several implications:

Firstly, it indicates the presence of a knowledgeable and experienced academic community at the University of Johannesburg, particularly in the area of ChatGPT. This suggests that the institution is actively involved in research and development in this field, potentially making it a hub for cutting-edge advancements and innovation.

Secondly, understanding the network characteristics of published academic papers can shed light on the collaborative efforts and partnerships between different institutions and researchers. In this case, it may reveal potential collaborations between the University of Johannesburg and other academic institutions or industry experts in the field of ChatGPT. This can have implications for knowledge sharing, interdisciplinary research, and the overall advancement of the field.

Thirdly, the identification of a specific active author as a professor at the University of Johannesburg can also have implications for the reputation and credibility of the institution. If the author is well-regarded and recognized in the academic community, it reflects positively on the university's academic standing and expertise in the field of ChatGPT. This can attract more researchers, students, and funding opportunities, further enhancing the university's reputation and potential for future advancements.

It is crucial to recognize that this research's implications extend beyond the immediate domain of ChatGPT. The insights gained from analyzing network characteristics can have broader applicability in shaping academic policies and strategies to promote healthy competition, knowledge dissemination, and continuous academic progress. By utilizing these valuable insights, stakeholders in the academic community can work collaboratively to create an environment that nurtures academic excellence, fosters innovation, and guards against potential monopolistic tendencies that could hinder the advancement of knowledge and education.

6. Conclusion

In conclusion, this paper analyzes ChatGPT academia, scholars' sentiments, and co-authorship networks. The research topics cover diverse themes, including benefits, challenges, and applications in various fields. Scholars generally exhibit positive attitudes toward ChatGPT, but concerns about potential risks and limitations are evident. The analysis of co-authorship networks highlights the need to promote diverse collaborations to avoid education monopolies. Overall, this study offers valuable insights for responsible AI integration and academic progress.

6.1. Ethical considerations

The ethical considerations pertaining to this research involve two potentially contentious behaviors that warrant careful examination.

Firstly, the mention of the specific actor, while contributing to the scrutiny of academic monopoly, raises ethical concerns related to confidentiality and privacy. Identifying individual actors without their explicit consent may compromise their privacy rights and could lead to unintended consequences, such as reputational damage or breach of academic integrity. To address this concern, future assessments of academic monopolies should be conducted while preserving the anonymity of specific actors to ensure compliance with ethical standards and safeguard individual privacy.

Secondly, the research's focus on Digital Object Identifiers (DOIs) and abstracts as a means to mitigate copyright concerns introduces ethical implications surrounding intellectual property rights. While using DOIs and abstracts can be a practical approach to avoid potential copyright infringement, it is essential to balance this with the need for proper attribution and recognition of authorship.

6.2. Future researches

ChatGPT holds promising prospects for the future, particularly in fields like medicine and business, where significant advancements are anticipated. For future scholars, we recommend focusing on the third area identified in our literature review, which involves examining ChatGPT's specific applications in these domains and the transformative effects it brings forth. By delving deeper into these areas, researchers can uncover novel insights and contribute to the continued development of ChatGPT's practical applications, fostering innovation and progress in diverse industries.

Acknowledgment

References

[1] Leggatt, J. (2023, July 31). What is chatgpt? A review of the AI in its own words. Forbes. https://www.forbes.com/advisor/business/software/what-is-chatgpt/

[2] Donald Nielsen September 26, 2022 Education. (2023, May 18). The monopolies of Education. Discovery Institute. https://www.discovery.org/a/the-monopolies-of-education/

[3] Liu, Y.-H., Han, T., Ma, S., Zhang, J.-Y., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., & Ge, B. (2023, May 11). Summary of ChatGPT/GPT-4 Research and Perspective Towards the Future of Large Language Models. https://www.researchgate.net/publication/369791983_Summary_of_ChatGPTGPT-4_Research_and_Perspective_Towards_the_Future_of_Large_Language_Models

[4] Ray, P. P. (2023, April 14). CHATGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems. https://www.sciencedirect.com/science/article/pii/S266734522300024X

[5] Pan, J., Lyu, H., & Luo, J. (2023, July 11). Excitements and Concerns in the Post-ChatGPT Era: Deciphering Public Perception of AI through Social Media Analysis. https://arxiv.org/pdf/2307.05809.pdf

[6] Biswas, S. S. (2023, March 1). Potential use of chat GPT in global warming - annals of biomedical engineering. SpringerLink. https://link.springer.com/article/10.1007/s10439-023-03171-8

[7] Burnham, J. F. (2006, March 8). Scopus database: A review - biomedical digital libraries. BioMed Central. https://doi.org/10.1186/1742-5581-3-1

[8] Person, Itay, & Abuhav. (2018, May 23). ISO 13485:2016: A Complete Guide to Quality Management in the medical. Taylor & Francis. https://www.taylorfrancis.com/books/mono/10.1201/9781351000796/iso-13485-2016-itay-abuhav

[9] Canché, M. S. G. (2023, April 1). Machine driven classification of open-ended responses (MDCOR): An analytic framework and no-code, free software application to classify longitudinal and cross-sectional text responses in survey and social media research. ScienceDirect. https://doi.org/10.1016/j.eswa.2022.119265

[10] Zhang, T. (2023). CHATGPT CIS SOCIOGRAM. RPubs. https://rpubs.com/Tongwei-Zhang/GPTSOCIOGRAM

[11] Saif, M. M., & Turney, P. (2011, July 10). NRC Word-Emotion Association Lexicon. NRC Emotion Lexicon. https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

[12] Andreotta, M., Nugroho, R., Hurlstone, M. J., Boschetti, F., Farrell, S., Walker, I., & Paris, C. (2019, April 2). Analyzing social media data: A mixed-methods framework combining computational and qualitative text analysis - behavior research methods. SpringerLink. https://link.springer.com/article/10.3758/s13428-019-01202-8

[13] Kraus, S. (n.d.). Sascha Kraus - Google Scholar Citations. Google Scholar. https://scholar.google.com/citations?user=SCyTFaYAAAAJ&hl=en

[14] George, A. S., George, A. S. H., & Martin, A. S. G. (2023, February 15). A review of Chatgpt Ai's impact on several business sectors. Zenodo. https://zenodo.org/record/7644359

[15] Liu, J., Wang, C., & Liu, S. (2023). Utility of chatgpt in clinical practice. Journal of Medical Internet Research, 25. https://doi.org/10.2196/48568

[16] 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

[17] Baidoo-Anu, D., & Owusu Ansah, L. (2023). *Education in the era of Generative Artificial Intelligence (AI): Understanding the potential benefits of CHATGPT in promoting teaching and learning. SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.4337484

[18] Fatani, B. (2023). *Chatgpt for future medical and Dental Research. Cureus.* https://doi.org/10.7759/cureus.37285

[19] Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). *Chatting and cheating: Ensuring academic integrity in the era of chatgpt. Innovations in Education and Teaching International,* 1–12. https://doi.org/10.1080/14703297.2023.2190148

[20] Li, Z. (2023, April 21). *The Dark Side of CHATGPT: Legal and ethical challenges from stochastic parrots and hallucination. arXiv.org.* https://arxiv.org/abs/2304.14347

[21] Wang, B., Yang, B., Shan, S., & Chen, H. (2019). *Detecting hot topics from academic big data. IEEE Access, 7,* 185916–185927. https://doi.org/10.1109/access.2019.2960285

[22] Rahman, Md. M., & Watanobe, Y. (2023a). *Chatgpt for Education and research: Opportunities, threats, and strategies. Applied Sciences, 13(9),* 5783. https://doi.org/10.3390/app13095783

[23] Oviedo-Trespalacios, O., Peden, A. E., Cole-Hunter, T., Costantini, A., Haghani, M., Rod, J. E., Kelly, S., Torkamaan, H., Tariq, A., David Albert Newton, J., Gallagher, T., Steinert, S., Filtness, A. J., & Reniers, G. (2023). *The risks of using CHATGPT to obtain common safety-related information and advice. Safety Science, 167,* 106244. https://doi.org/10.1016/j.ssci.2023.106244

[24] Coulson, A. J. (1996). *Markets versus monopolies in Education. Education Policy Analysis Archives, 4,* 9. https://doi.org/10.14507/epaa.v4n9.1996