



Saarth E-Journal

Saarth E-Journal of Research

E-mail: sarthejournal@gmail.com
www.sarthejournal.com

ISSN NO: 2395-339X
Peer Reviewed
Vol.9, Issue.3 No.07

Impact Factor: 6.89
Quarterly
July to Sept. - 2024

MAPPING THE KNOWLEDGE ROADMAP OF CHATGPT: A SCIENTOMETRIC ANALYSIS

*Ms. Hiral Patel, **Dr. Vipul Banker

ABSTRACT

The primary aim of the research is to investigate scientometrics parameters of scholarly publications in the field of ChatGPT for the year 2022 & 2023. The Scopus database was employed to retrieve the information regarding scientific publications. The search string “ChatGPT” were applied to find literatures published during the said period. The first scientific article on ChatGPT came across the scopus database in 2022. Total of 355 scholarly publications were found. Among them, 4 documents were deleted from the dataset as they appeared to be incorrect data. After data mining process the 351 scholarly publications retrieved for analysis and network of visualization using scientometric parameters along with MS-Excel and VOSviewer tools. There has been a tremendous growth of research in a single year. The total citations 29 and 648 have been received in the year 2022 & 2023 respectively. The CAGR for the scholarly publications of ChatGPT were recorded 12.21 in the year 2023. The highest publications were observed in the subject area ‘medicine, social science, computer science’ for the ChatGPT research. In ChatGPT research, article-type documents have been seen published more than all other types of documents. Artificial Intelligence keyword has been appeared in the research productivity. The highest number of papers published by Wang, F.Y. (7 papers), followed by Wu, H. (5 papers). The majority of scientific literatures in this research development came from the USA. The visualization network was exhibited in terms of co-occurrences, co-citation analysis, co-authorship analysis using VOSviewer.

KEY WORDS: ChatGPT, Artificial Intelligence, Scientometrics, VOSviewer, Authorship Pattern, Scientometric Parameters.

(*Ms. Hiral Patel, Research Scholar, **Dr. Vipul Banker, Research Guide, Surendranagar University, Surendranagar.)

INTRODUCTION

OpenAI is a research center established in the year of 2015 (Brockman, et al., 2016). OpenAI has grown swiftly and became most well-known in technology. The artificial intelligence company startup has created ChatGPT which has become its most prominent invention (Hughes, 2023). ChatGPT is a language model which introduced by OpenAI in November 2022 (Hetler, 2023). GPT stand for Generative Pre-Trained Transformer. The Training for the model is intended to make it possible to produce text that is logical, pertinent, and human-like (Gent, 2023). It is a natural language processing tool executed by artificial intelligence (AI). It performs like human discussions and is much more compatible with the chatbot or it is an artificial chatbot that uses machine language processing to construct & understanding conversations (Ortiz, 2023). The language model is capable of answering queries as well as composing a variety of written content, such as essays, code, articles, emails, and social media postings. ChatGPT is equivalent to the automated chat services that exist on consumer support websites where people can send request or ask questions regarding their queries and get answer from chatGPT in a real time (Hetler, 2023).

Use of ChatGpt:

The generative model is used in a particular manner to compose music, draft emails, presentations, write cover letter, prepare resume, explain topics, solve math difficulties, provide tips for relationships, generate content in several languages etc. In a nutshell, it has the ability to perform multiple kinds of language-based tasks (*ChatGPT and Its Impact on Society*, 2023).

How to login & access ChatGPT:

The process of ChatGPT access is and also exhibited in figure 1. In figure 1, the steps of procedure are ordered from left to right.

- Enter URL in Browser: <https://openai.com/>

- Click the sign up button located the top right corner if you haven't already logged in.
- Create an account or sign in using your Google, Microsoft, or Apple credentials.
- Among the three options, select ChatGPT.
- For conversation, the ChatGPT window is now open.
- Put a query/Questions which we want to know in the "send a message" box.

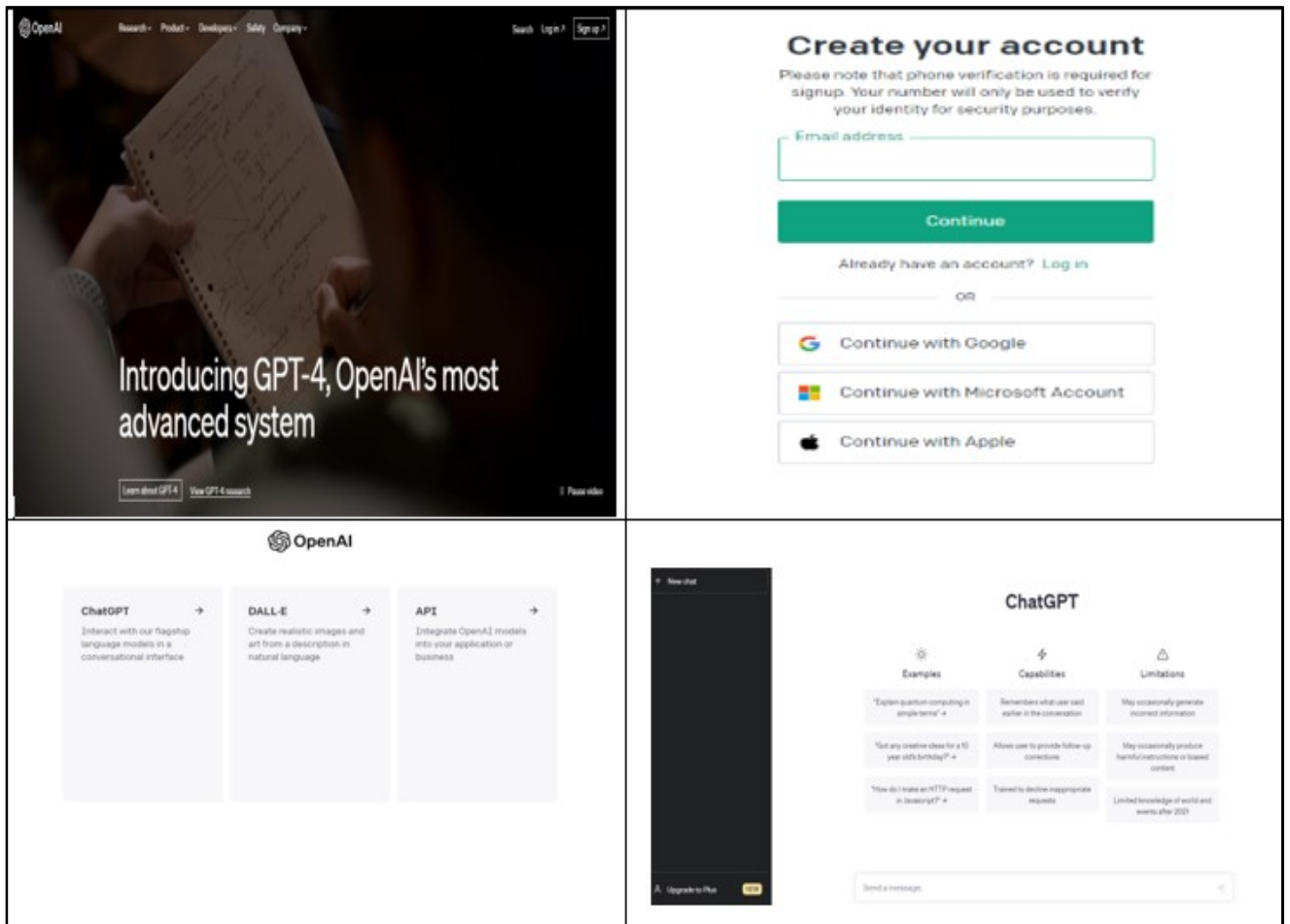


Figure 1: Process for ChatGPT Login

Examples of ChatGPT Creations:

In this part, every question mentioned with a "Q." was created by a human interviewer, and every answer mentioned with a "A." was constructed totally by ChatGPT.

Q. write article on chatGPT

A. ChatGPT is an AI-powered conversational agent that has been trained using the latest advances in natural language processing. As a large language model based on the GPT-3.5 architecture, ChatGPT has the ability to understand and respond to human language in a way that is natural and intuitive.

One of the key strengths of ChatGPT is its ability to generate human-like responses to a wide range of queries and topics. Whether you're looking for information about a specific topic, or simply want to engage in casual conversation, ChatGPT is capable of providing engaging and informative responses.

But what sets ChatGPT apart from other conversational agents is its ability to learn and adapt over time. By analyzing the vast amounts of data, it processes on a daily basis, ChatGPT is able to continually improve its understanding of language and its ability to generate relevant and useful responses.

In addition to its language processing capabilities, ChatGPT also features a user-friendly interface that makes it easy to engage with. Whether you prefer to chat via text or voice, ChatGPT can accommodate your preferences and provide an experience that feels natural and intuitive.

But perhaps the most impressive thing about ChatGPT is its potential to revolutionize the way we interact with technology. With the ability to understand and respond to human language in a way that is natural and intuitive, ChatGPT has the potential to transform the way we search for information, communicate with one another, and even interact with the world around us.

For example, ChatGPT could be used to create virtual assistants that are capable of understanding and responding to complex queries in real-time. This could be particularly useful in industries like healthcare, where virtual assistants could be used to provide patients with personalized care and support.

Overall, ChatGPT represents a major leap forward in the field of natural language processing and AI. As it continues to evolve and improve, it has the potential to transform the way we interact with technology and each other, ushering in a new era of intuitive and engaging communication.

ChatGPT in Library regarding question:

Q. How can chatgpt useful in library?

A. ChatGPT can be a useful tool in a library in a number of ways:

1. Reference Assistance: ChatGPT can assist library users in finding relevant books, articles, and other resources by answering their questions about topics, authors, titles, and keywords. ChatGPT can also suggest additional sources that the user may not have considered.

2. Recommendations: ChatGPT can provide personalized reading recommendations to library users based on their interests, reading history, and preferences. This can help users discover new authors and books that they might not have otherwise known about.
3. Chatbot: ChatGPT can be used as a chatbot to provide instant assistance to library users who have questions or need help with basic tasks, such as renewing their library card, checking their account status, or finding the library's hours of operation.
4. Language Translation: ChatGPT can also be useful in libraries that serve multilingual populations by providing translation services in real-time. Library users can type their questions or requests in their preferred language, and ChatGPT can translate and respond in the user's preferred language.

Overall, ChatGPT can enhance the user experience and help libraries better serve their patrons by providing fast, accurate, and personalized assistance.

Scientometric analysis:

In accordance with the stated objectives of research productivity of ChatGPT, the outcomes of the study depend on scientometrics analysis using various parameters and mapping networks between articles which were published in scopus indexed journals.

Literature Review

The work of this paper is focusing on scientometric analysis of ontology research regarding construction. The data was collected from scopus for the period of 2007 – 2017 and retrieved total 199 records. As per their findings, the maximum numbers of literatures were found in the year 2011 & 2016. The results revealed that three countries U. S., U. K. and Canada has been carried out more than half publications (53%). In this study, the keywords were identified by cluster analysis (Zhong et al., 2019).

Artificial Intelligence is a relatively new technology. Research is increasing every day. In this paper, the data regarding geohazard research has been collected from web of science. The scientometric analysis tools such as co-authorship, cluster analysis and co-citation analysis have been applied to the 9226 Publications. The outcome demonstrates the tremendous growth of research after the year 2000, recorded in China, U. S. & Italy (Jiang et al., 2022).

the study was conducted to explore scientometric analysis in the field of plastic sand. Total 4512 publications were found in this field during the year 2001 to 2022 and data extracted from scopus database. The analysis performed for 4512 publications using VOSviewer and multiple scientometrics techniques. The subject areas were analysed and revealed that engineering, materials science and environmental science has most publications with 32.5%, 17.1% and 13.9% respectively. The largest numbers of documents were published under the keywords sand, compressive strength, and plastic. The minimum numbers of literatures by author has been Set at 10 and at that level, 36 authors satisfied this requirement. Wang is the most Prominent author with 30 shares. followed by wang 27, Zhang 22, and Li 21 shares. The Maximum numbers of citations received by Wu G (Iftikhar et al., 2023).

Scientists have tried to detect the research fronts of computer science through Scientometrics parameters. They evaluate leading institutions of country in the field of computer science. The topics deep learning, block chain, natural language processing, internet of things (IOT) became popular topics in this research (Xia et al., 2021).

Researchers has evaluated research trends on Social Support in Education for the period of 1977- 2020. The Scopus database was used to extract bibliographic records. The literatures have been visualized and analysed using VOSviewer, MS-Excel, ScientoPy, Citespace & Biblioshiny. The study assessed over all publications, author collaboration, most contributed countries & organizations, co-occurrences of keywords, etc., (Rashid et al., 2021).

Major Objectives

- ❖ To determine the current pillars of publications in the sphere of ChatGPT research.
- ❖ To identify annual growth rate (AGR) and Relative growth rate (RGR)
- ❖ To detect most productive organizations, countries and authors.
- ❖ To identify most preferred journals by ChatGPT scholars.
- ❖ To understand visualization networks such as co-occurrence of keywords, co-citations, co-authorship

Methodology

1.1 Method: The bibliographic details has been retrieved regarding scientific literatures of chatGPT. Furthermore, we looked at the conceptual, intellectual and social structures of the scientific literatures.

Conceptual: Co-word analysis through Co-occurrence of terms.

Intellectual: Co-citation networks (References co-citation, Journal co-citation)

Social: Collaboration network (Author & Country Collaborations).

1.2 Database Selection: The scopus is an abstract and citation database introduced by Elsevier in 2004 (Schotten et al., 2017). In this study, the scopus database utilized to collect bibliographic data for the period of 2022 to 2023. The ChatGPT has been initiated in 2022. The primary data was collected on May 10th, 2023.

1.3 Search Phrase: The following search query created and entered in search engine of scopus by operators.

Select article title, abstract, keywords

Applied search string: TITLE-ABS-KEY ("ChatGPT")

1.4 Data Extraction: A total of 355 scientific articles were obtained by applying the above-mentioned search string. 4 documents were identified as irrelevant data; therefore, they have been excluded from the datasets. The final 351 documents were considered for analysis and visualisation.

1.5 Data Analysis: The yielded documents were analysed using multiple scientometrics parameters to produce valuable data. Some fundamental operations performed by using Microsoft Excel. VOSviewer is a tool that assisted authors for data visualizations. The tool is helped in discovering, exploring & mapping co-occurrences of keywords, correlations between countries & authors, co-citation networks.

Scope and Coverage

Research Scholars examined scholarly papers in terms of "ChatGPT" that were published in journals, listed in Scopus index. This analysis was conducted for the years 2022–2023. The 'ChatGPT' is revealed in November 2022. therefore, the coverage period is 2022-2023.

Data Analysis and Interpretation

1.6 Annual Distribution of Publications:

Since the commencement of ChatGPT, it has gained an ample amount of attention. Figure 2 indicates the total number of scholarly publications, published in scopus indexed

journals. The first publications have been published in 2022 with 2 literatures. There has been a huge explosion in just one year in research activity on ChatGPT, and 349 works have been recorded in 2023.

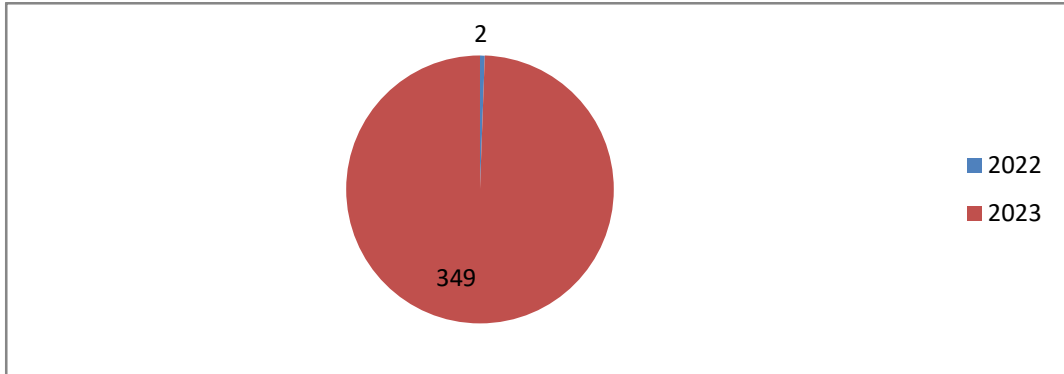


Figure 2: Annual Distribution of Publications

1.7 Research growth of ChatGPT:

Figure 3 shows research growth of chatGPT literatures. The annual growth rate (AGR) and compound growth rate (CAGR) mentioned in Table 2 is calculated based on the end value and the first value. It can be said from its output that the research trend is increasing significantly. The AGR & CAGR was recorded 17350 & 12.21 in 2023 respectively.

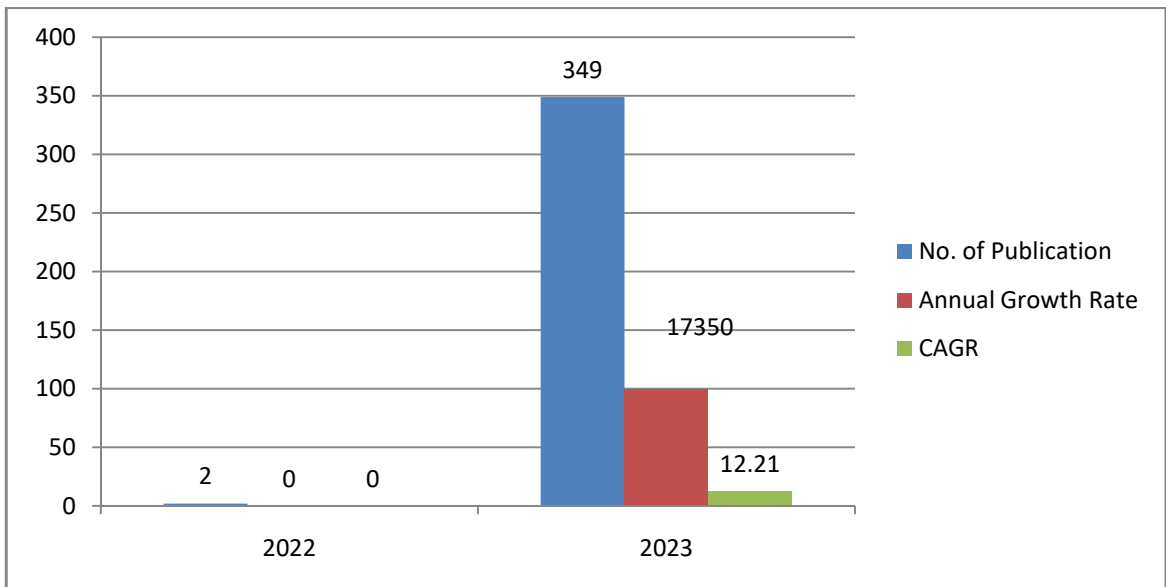


Figure 3: AGR and CAGR of Publications

1.8 Subject areas of research output:

The ChatGPT is a new invention, and it appears to be expanding quickly across many topics.

The field of medicine delivered the most research output with 149 publications out of 351, followed by the social sciences (99 papers) and computer science (60 papers). The overall covered areas are mentioned in figure 4.

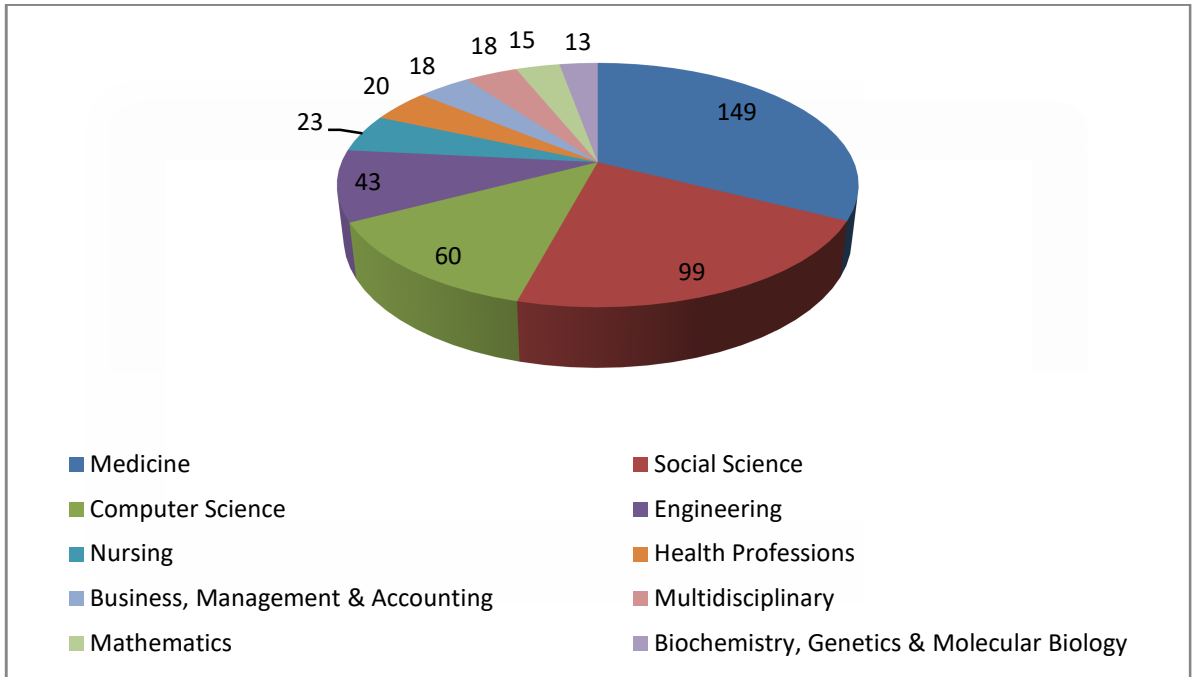


Figure 4: Subject areas of research output

1.9 Types of documents for ChatGPT research:

Figure 5 contained information regarding various types of documents which involves in ChatGPT research activity for the said period. The large numbers of scholarly literatures in the area of ChatGPT have been published in Article with 120 papers, 33.80%, followed by Editorial with 84 papers, 23.66%, Letter with 63 papers, 17.75%. The other types of documents have been listed in figure 5 with their research works.

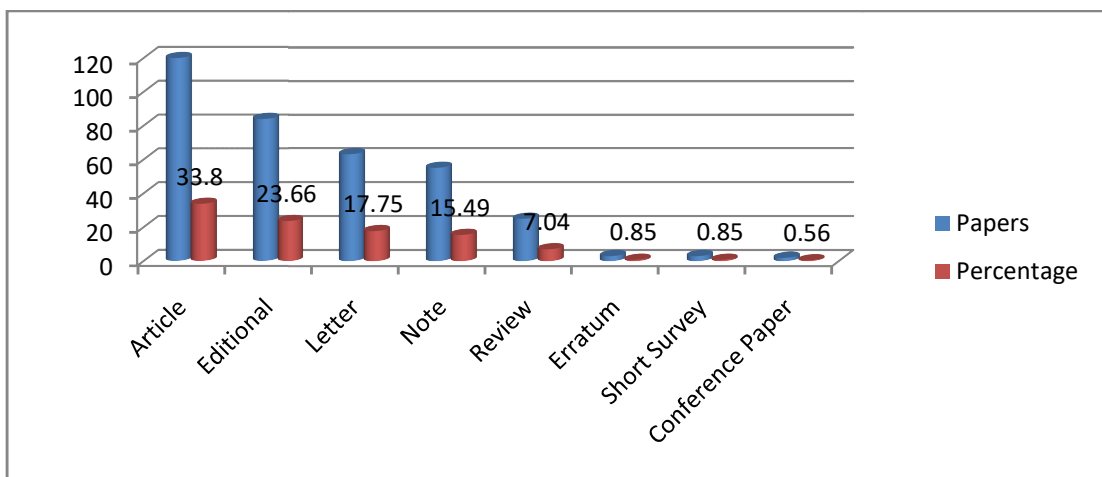


Figure 5: Document-wise Distribution

1.10 Most prominent countries:

Figure 6 summarizes the most contributed countries on ChatGPT research activity in context of publications. The United States has delivered the maximum numbers of literatures with 105 papers (29.58%), followed by United Kingdom with 31 papers (8.73%), China has involved in this research with 24 papers (6.76%). The statistical data of other leading countries described in figure 6.

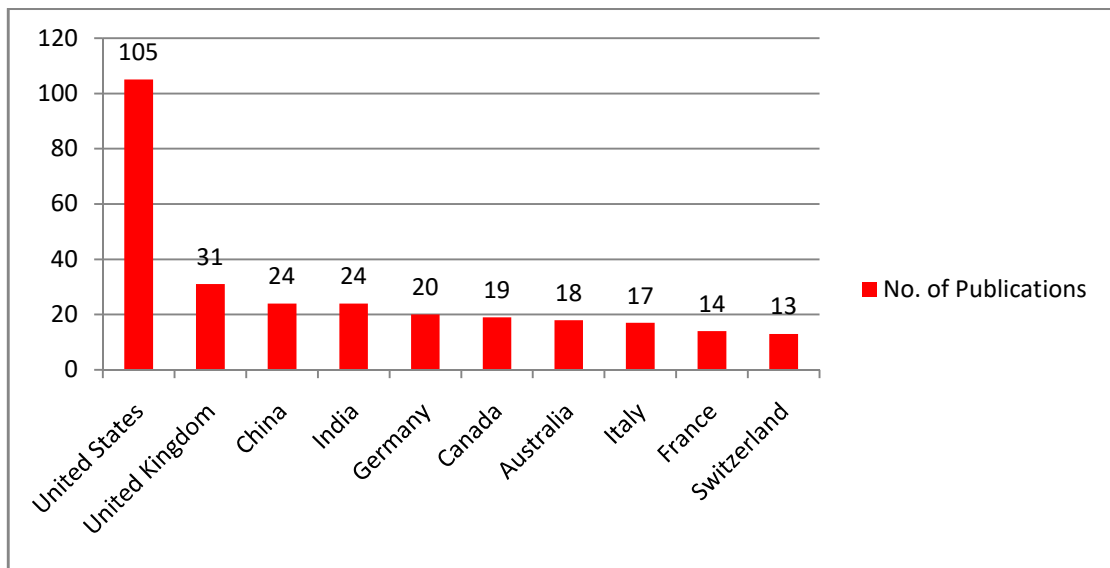


Figure 6: Most prominent countries

1.11 Most contributed organizations:

Figure 7 depicted the numerical data regarding papers, percentage as per whole research and percentage as per top ten organizations for the most contributed organizations in ChatGPT research productivity. According to the presented data in figure 7, Chinese Academy of Sciences has produced the most literature with 7 papers, 1.97% as per the whole research, then AP-HP Assistance Publique - Hopitaux de Paris and University of Chinese Academy of Sciences have been involved in a similar proportion with 6 papers and 1.69% shares.

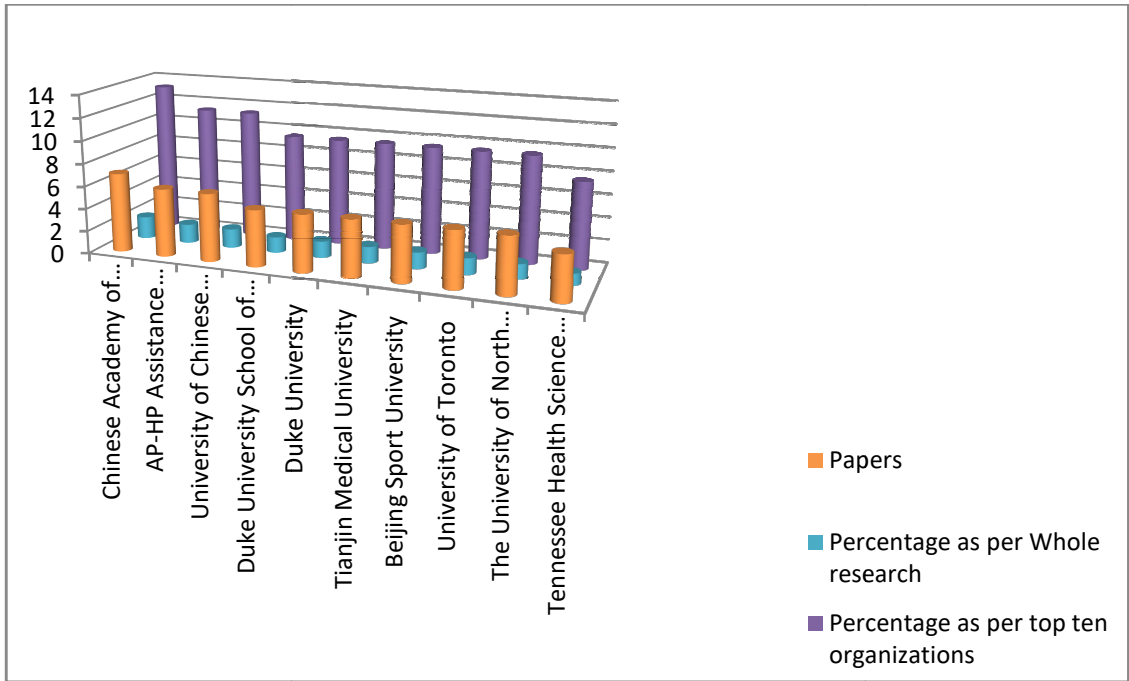


Figure 7: Most contributed organizations

1.12 Co-authorship analysis in relation of authors:

The author collaboration network is outlined in figure 8. The authors represented by nodes. The dimension of a node is related to the quantity of papers the author has written. According to figure 8, there were 2 clusters observed, each in a different colour. The total network link 58, total link strength 144 were depicted. An additional examination of co-authorship collaboration determined that "He, Y. and Wu, H." acquired total 6 links and 21 total link strength with greatest collaboration followed by Cheng, K., Gu. S. with total link strength 19, 17 respectively. Figure 9 shows most prolific authors in context of ChatGPT research activity.

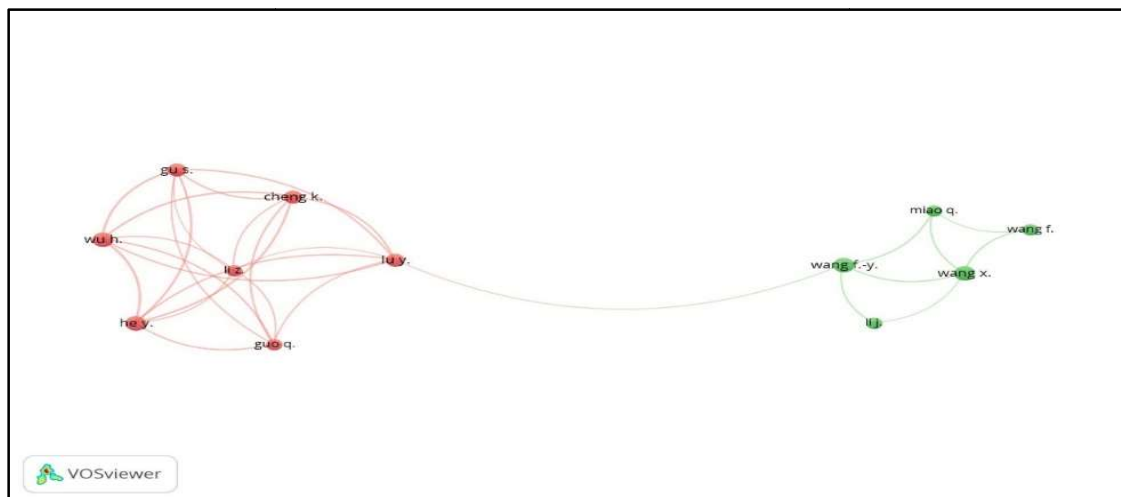


Figure 8: Co-authorship analysis in relation of authors

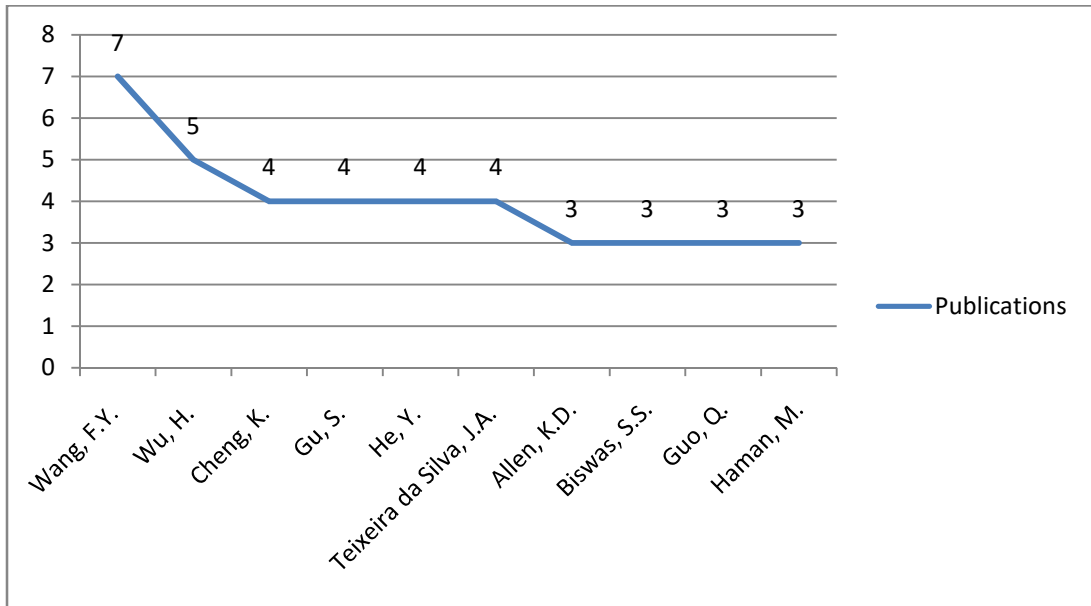


Figure 9: Top ten authors as per documents

1.13 Co-authorship analysis in relation of countries:

The co-authorship analysis related to countries demonstrates the collaborative efforts between countries in this area. The bigger nodes reflect the most contributed nations in the field of ChatGPT. The collaboration between countries is represented by the thicker lines of nodes. As shown in figure 10, the 11 most contributed countries in the field of ChatGPT from 4 clusters which highlighted by different colours. The United States has the highest link strength (28), documents (105), and citations (191) followed by United Kingdom with 13 total link strength, 30 documents, 74 citations. China has got third position with 11 total link strength, 28 documents, 29 citations. Furthermore, Switzerland has got 31 citations. Furthermore, Switzerland has also received 31 citations for the 12 documents.

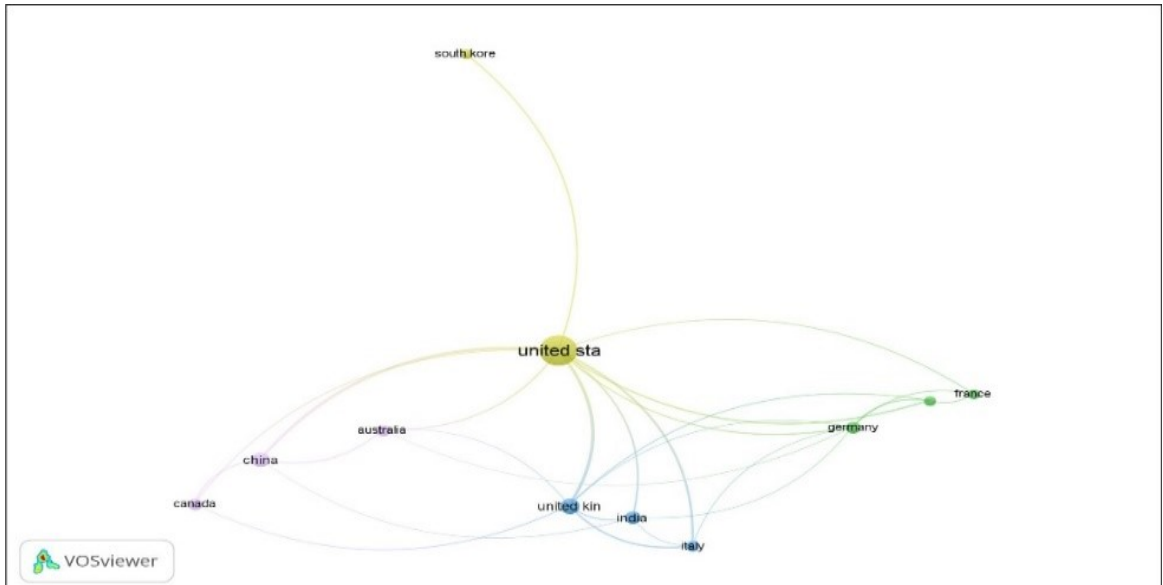


Figure 10: Co-authorship of countries

1.14 Co-occurrences of keyword based on author keyword:

The map of keyword co-occurrences for ChatGPT research was designed to evaluate and understand the research activity in this area. As per figure 11, Total 23 keywords were met the threshold based on minimum number of occurrences of a keyword set at 5 thresholds.

The stronger co-occurrence relationships between terms were represented by the thicker lines. The most frequency of occurrences of keywords were represented by larger circles. The term “ChatGPT” has the maximum total link strength 234 with 126 occurrences followed by artificial intelligence with total link strength 192 and 92 occurrences.

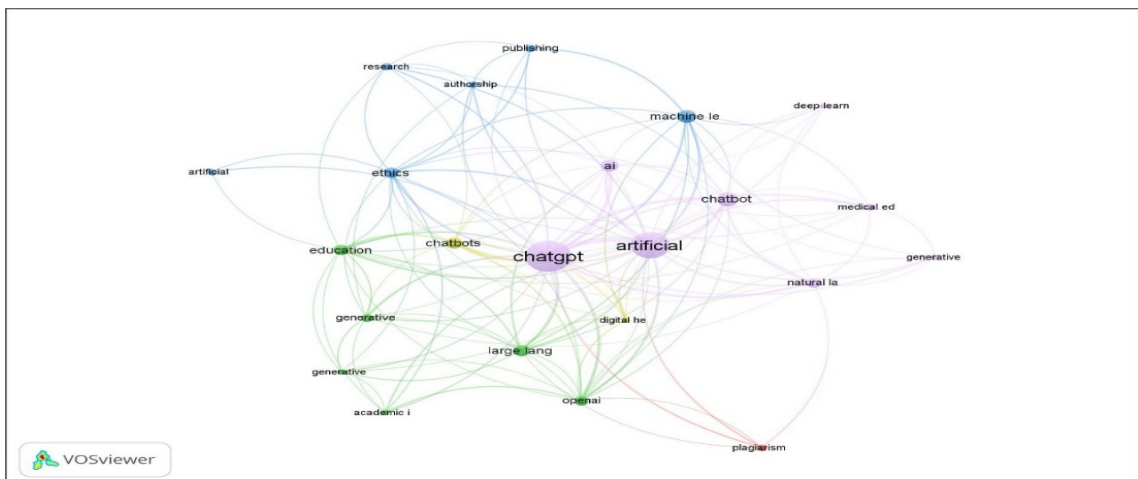


Figure 11: Frequency of keywords occurrences

1.15 Visualization network of Citation analysis:

The map of author's citation networks is illustrated in figure 12. Authors denoted by nodes. The node size represents weightage of author's citation frequency. Total 4 clusters with the total link strength 324 and 100 links were observed in the visualization network. He Y. and Wu H. have 46 links.

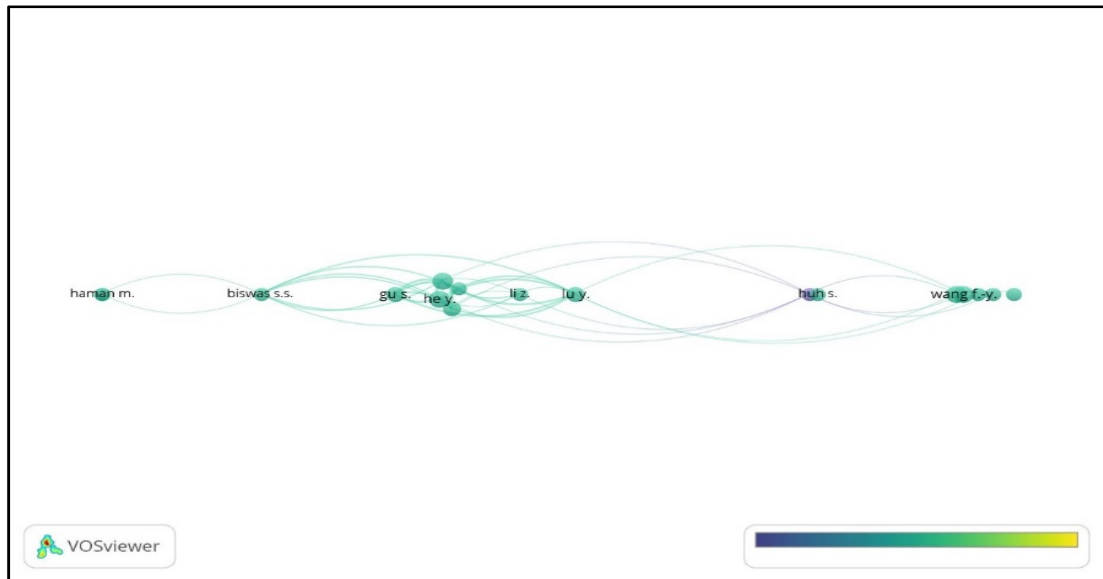


Figure 12: Citation analysis networks

1.16 Co-citation analysis – Source Title

Figure 13 illustrates co-citation networks of sources. The analysis revealed the citation links between the other sources. Total 5 clusters, 274 links, 6250 total link strength were observed in the context of ChatGPT research trends. The clusters, weight of links and link strength that emphasis the importance of citations of source titles in the network visualization. According to map network, nature has the highest links 21, total link strength 1390, &citations 246. So it can be decided that it has the strongest network with other sources in context of ChatGPT research.

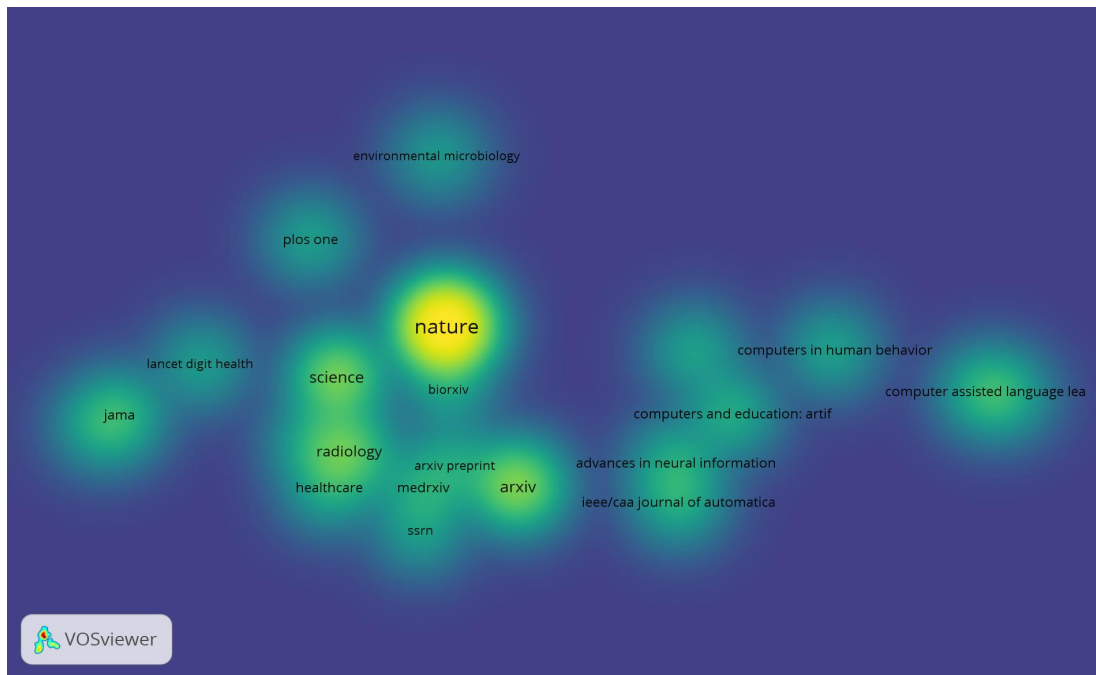


Figure 13: Co-citation analysis in context of source title

Findings:

The findings from various studies and analyses provide valuable information regarding the capabilities, limitations, and areas for future study in the context of ChatGPT. One finding is that ChatGPT demonstrates impressive proficiency in generating coherent and contextually relevant responses in conversation. The analysis of the data reveals a remarkable growth trend in the research activity of ChatGPT over the span of one year. Findings indicate an astounding increase of 17350% (from 2 to 349 papers) in ChatGPT output. In terms of subject areas, findings indicate that ChatGPT research is rapidly growing with 42.45% in medicine domain. The findings reveal that the United States United Kingdom and China have emerged as the most prominent country with 29.58%, 8.73%, 6.76% respectively in ChatGPT research output. Chinese Academy of Sciences is the most effective organisation, contributing 1.97% of all research effort.

Conclusion:

The present investigation focuses on scientometric analysis and VOSviewer software tool provide visualization network of the ChatGPT research. The field of ChatGPT is evolving rapidly in various domains including medical, social science,

computer science, engineering, etc. The USA is the most prominent country; the UK and China also have significant places across the globe in research productivity. Although ChatGPT operated by AI is expanding into research disciplines, a scientometric analysis emphasising the logical base, research subject development, and research areas needs to be developed. This has laid down the rational basis and provided directions for the emergence of research subjects and hotspots.

This acknowledges existing research growth in ChatGPT while likewise providing opportunities for further study. Further research can be conducted using the WoS (Web of Science) database, Dimensions, and Google Scholar, which may provide greater insight in the context of ChatGPT research productivity.

References:

1. Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). Openai gym. arXiv. <https://doi.org/10.48550/arXiv.1606.01540>
2. ChatGPT and its Impact on Society. (2023, February 13). Times of India Blog. <https://timesofindia.indiatimes.com/readersblog/marketing-savvy/chatgpt-and-its-impact-on-society-50445/>
3. Gent, E. (2023, March 29). Hello, ChatGPT—Please Explain Yourself! *IEEE Spectrum*. <https://spectrum.ieee.org/chatbot-chatgpt-interview>
4. Gomis, M. K. S., Oladinrin, O. T., Saini, M., Pathirage, C., & Arif, M. (2023). A scientometric analysis of global scientific literature on learning resources in higher education. *Heliyon*, 9(4), e15438. <https://doi.org/10.1016/j.heliyon.2023.e15438>
5. Hetler, A. (2023). ChatGPT. *WhatIs.com*. <https://www.techtarget.com/whatis/definition/ChatGPT>
6. Hughes, A. (2023). ChatGPT: Everything you need to know about OpenAI's GPT-4 tool. *www.sciencefocus.com*. <https://www.sciencefocus.com/future-technology/gpt-3/>
7. Iftikhar, B., Alih, S. C., Vafaei, M., Alrowais, R., Bashir, M. T., Khalil, A., Rizwan, M., Javed, M. F., Rehman, M. F., & Mohamed, A. (2023). A scientometric analysis approach on the plastic sand. *Heliyon*, 9(3), e14457. <https://doi.org/10.1016/j.heliyon.2023.e14457>
8. Jiang, S., Ma, J., Liu, Z., & Guo, H. (2022). Scientometric Analysis of Artificial Intelligence (AI) for Geohazard Research. *Sensors*, 22(20), 7814. <https://doi.org/10.3390/s22207814>
9. Ortiz, S. (2023). What is ChatGPT and why does it matter? Here's what you need to know. *ZDNET*. <https://www.zdnet.com/article/what-is-chatgpt-and-why-does-it-matter-heres-everything-you-need-to-know/>
10. Rashid, S., Rehman, S. U., Ashiq, M., & Khattak, A. (2021). A Scientometric Analysis of Forty-Three Years of Research in Social Support in Education (1977–2020). *Education Sciences*, 11(4), 149. <https://doi.org/10.3390/educsci11040149>
11. Schotten, M., Aisati, M. E., Meester, W. J. N., Steinginga, S., & Ross, C. (2017). A Brief History of Scopus: The World's Largest Abstract and Citation Database of Scientific Literature. In *Auerbach Publications eBooks* (pp. 31–58). <https://doi.org/10.1201/9781315155890-3>

12. Xia, W., Jiang, Y., Zhu, W., Zhang, S., & Li, T. (2021). Research Fronts of Computer Science: A Scientometric Analysis. *Journal of Scientometric Research*, 10(1), 18–26. <https://doi.org/10.5530/jscires.10.1.3>
13. Zhong, B., Wu, H., Li, H., Sepasgozar, S. M. E., Luo, H., & He, L. (2019). A scientometric analysis and critical review of construction related ontology research. *Automation in Construction*, 101, 17–31. <https://doi.org/10.1016/j.autcon.2018.12.013>