AIMS Mathematics

AIMS Mathematics, 9(7): 19737–19755.
DOI: 10.3934/math.2024963
Received: 08 February 2024
Revised: 05 June 2024
Accepted: 11 June 2024
Published: 17 June 2024

https://www.aimspress.com/journal/Math

Research article

A study of the impact of scientific collaboration on the application of Large Language Model

Suyan Tan*, Yilin Guo

School of Foreign Studies, Guangzhou University, Guangdong 510006, China

* **Correspondence**: Email: sfstansuyan@gzhu.edu.cn.

Abstract: The study of Large Language Models (LLMs), as an interdisciplinary discipline involving multiple fields such as computer science, artificial intelligence, and linguistics, has diverse collaborations within its field. In this study, papers related to LLMs in the SSCI and SCI sub-collections of the Web of Science core database from January 2020 to April 2024 are selected, and a mixed linear regression model is used to assess the impact of scientific collaborations on the application of LLMs. On this basis, the paper further considers factors such as financial support and dominant countries to deeply explore the heterogeneous impact of scientific collaborations on the application of LLMs. The findings show that (1) excessive involvement of academic institutions limits the research and application of LLMs, and the number of authors does not have a significant effect on the application of LLMs; (2) with or without financial support, the role played by scientific collaborations in the application of LLMs does not significantly change; and (3) differences in the dominant countries of scientific collaborations have a slightly heterogeneous effect on the role of LLMs applications, which are mainly reflected in the number of collaborators.

Key words: scientific collaboration; Large Language Model application; financial support; dominant countries; heterogeneity
Mathematics Subject Classification: 94B, 35

1. Introduction

Large Language Models (LLMs) have emerged as a hotspot in the academic field due to their remarkable learning capabilities and versatility. These models are rooted in deep learning, harness vast amounts of textual data to comprehend human language, and generate text with natural linguistic styles [1]. Presently, research achievements in the field of LLMs have been applied in various domains, such as in medicine and in bioinformatics areas. In the medical domain, the utilization of LLMs centers on tasks such as biomedical automatic question answering, text summarization, search engine optimization, and information extraction. Med-PaLM2, one of Google's medical models, has an accuracy rate comparable to that of a human physician. Med-PaLM2 is the first model to achieve expert test-taker level performance on the MedMCQA dataset of the United States Medical Licensing Examination (USMLE), with an accuracy rate of more than 85%, and is also the first AI system to achieve a passing score of 72.3% on the MedMCQA dataset, which includes questions from the Indian medical exams AIIMS and NEET, achieving unprecedented accuracy in the clinic [2]. The digitization of healthcare and documentation of electronic health records is also becoming a trend [3]. In the realm of bioinformatics, LLMs demonstrate their excellent performance in protein function prediction [4]. Notably, researchers have employed models, such as PDB Bind and STCRDAB, to facilitate the efficient prediction of functional associations among compounds, genes, and proteins. Such advancements enhance promises for enhancing disease diagnosis and expediting drug discovery [5]. Through the characterization of the literature related to Large Language Model (LLM) research, it is found that most of the keywords with high co-occurrence frequency focus on the field of education, such as "education", "artificial intelligence", "English, children", "Large Language Model", "generative AI", etc. This reflects the high degree of integration between LLMs and the field of education, and also indicates that the field of education is a popular application area of LLMs. The booming development of LLMs has led to the era of human-computer collaborative development and educational symbiosis. ChatGPT as a virtual teaching aid can provide students with personalized and instant knowledge, as well as interactive learning and assessment [6]. The use of ChatGPT as a problem-solving tool in the field of pedagogy can help address the shortcomings of traditional classrooms that focus on book-based basics, provide students with more hands-on opportunities, develop critical thinking, improve problem-solving skills, and promote deeper understanding of core subjects [7].

In order to enhance international competitiveness and productivity, various countries or regions have issued relevant policy or documents to promote the development of LLMs in various fields. Taking the development of LLMs in the field of education as an example, countries and regions have promoted its further development and application by adopting policy guidelines. There are at least 20 relevant policy guidelines and reports have been issued in 2023. For example, in May, the U.S. Department of Education's report on "Artificial Intelligence and the Future of Teaching and Learning" pointed out that AI brought important opportunities to address education policy priorities, but at the same time, attention must be paid to possible accidents and risks. In July, Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT) Bureau of Elementary and Secondary Education issued the first "Interim Guidelines for the Utilization of Generative AI in Elementary and Secondary Education" and pointed the way for the application of generative AI in the education field. In September, UNESCO issued the world's first generative AI-related guideline document, the "Guidelines for Generative Artificial Intelligence in Education and Research", which is aimed at enabling better integration of generative AI into education. The 12th China Intelligent Industry Summit held in the same year released a white paper on "Educational Applications of LLMs", which is intended to systematically promote the application of LLMs in the field of education. In November, Nature published an article discussing how to seize the opportunity brought by the transformative tool of

LLMs to reinvent the education industry. From the above, it is clear that different countries are actively engaging in the application of LLMs in education, which is not only profoundly reshaping the way human beings interact with technology, but also is gradually integrating it into people's daily life, work, and learning [8].

2. Literature review

Currently, the focus of academic research on LLM applications has been diversified, from focus on education to the proliferation of natural language processing, medicine, and other fields. With the help of the Citespace software, this paper analyzed the literature from the core collection of Web of Science from January 2020 to April 2024 to learn about the evolution path, research hotspots, and scientific collaboration in the field of LLMs.

The development of LLMs has gone through a developmental period (2020-2021) and an explosive period (2022-2024), with only 895 publications in the developmental period and 1575 in the explosive period, which was partly due to the fact that OpenAI optimized their model for question-and-answer dialogues in 2021, and this set off a small upsurge in the research of LLMs. Further optimization in 2022 resulted in the language model ChatGPT, and OpenAI opened it to the public for trial in November 2022, which pushed research on LLMs into an explosive period. The focus of scholars at different development stages also changed accordingly. Scholars in the developmental period mainly focused on the combination of AI and the education field, mainly paid attention to the research object of educators, and gradually the technical problems of natural language processing and machine learning were also included in the scope of research on LLMs. Schillinger et al. [9] mentioned in their findings that researchers could utilize the techniques of natural language processing and machine learning to build novel, scalable, and effective health literacy profiles that can be used to identify patients with limited health literacy, thus improving the treatment efficiency of healthcare. There was an explosion of scholars who devoted their attention to exploring the application of LLMs in different disciplinary areas. Indran et al. [10] mentioned in their study that formulating quality assessment questions in medical education was a critical but time-consuming, expertise-driven endeavor that requires innovative solutions, and LLMs such as ChatGPT offer a novel but not yet fully explored avenue for such innovation. Applying them to medical problems can improve educators' productivity and enable students to learn more about themselves. Meanwhile, Divito et al. [11] focused on the immense potential of LLMs or generative AI techniques in the field of education, such as the important role played by ChatGPT in transforming the education system, as well as enhancing the learning outcomes of students. In addition to this, LLMs also play an important role in the field of finance. Different machine learning models can predict stock price fluctuations to some extent. The highest average annualized return of 29.57% was obtained during a simulated trading experiment to predict the performance of the machine learning-based model framework [12]. To summarize, the development of the application of LLMs as a whole presents a multidimensional research perspective and the phenomenon of diversified and interdisciplinary research.

In order to further understand the research hotspots of LLMs, literature data was clustered and analyzed[1] (Figure 1), and the main keywords under each cluster were summarized (Table 1). Four

[1] Cluster analysis of the literature data in Citespace can get the keyword clustering map, and in the clustering map, if the Q value is greater than 0.3, the S value is greater than 0.5, indicating that the structure and homogeneity of the clustering view is reasonable, and the overall drawing effect is better.

major research hotspots were determined, as follows:

(1) Application of machine learning in LLMs. Machine learning is a class group with the largest cluster size in the keyword clustering atlas, and this class group contains the largest number of keywords, including "model", "machine learning", "natural language processing", "aggregation operators", and so on. Natural language processing is an important branch in the field of artificial intelligence, which aims to allow computers to understand, generate, and process human language, and the main tasks include speech recognition, machine translation, sentiment analysis, text summarization, and question and answer systems. New natural language processing techniques and LLMs will improve patient understanding of reports, and AI can simplify imaging reports without causing major disruptions, thus making it easier and faster for patients to understand their individual imaging reports [13]. Natural language processing is a branch of artificial intelligence just like machine learning and deep learning, and because it deals with natural language, it is also said to be the intersection of artificial intelligence and linguistics. The emergence of artificial intelligence provides many opportunities for human beings, the application scenarios of artificial intelligence are also very wide, and scholars' research on natural language processing for machine learning can make greater use of modern tools to process data efficiently as well as to improve the total factor productivity [14]. For example, people use search engines every day, and they generate a large amount of information which needs to be categorized and extracted to achieve efficiency in the dissemination of information [15], as well as accuracy in the prediction of trend [16]. Scholars' interest in natural language processing centers on its application of how to guide computers to process large-scale data with more efficient algorithms.

(2) Large Language Models in healthcare. There are two clusters in this hotspot, including Large Language Model and healthcare, and the co-occurring keywords related to these themes are "artificial intelligence", "Large Language Model", "health", "healthcare", "medical education", and so on. Existing research on LLMs in medical areas focus on biomedical automated question and answer, medical decision making, electronic records, etc. The application of LLMs in medical areas is still in the exploratory stage, and needs to be further improved in order to solve complex language comprehension problems in a more effective way and to generate simpler and more accurate solutions [17]. For example, in order to answer biomedical questions well, models first need to have knowledge of biomedicine to target the correct diagnosis of a patient. Second, they need to be able to read a very long medical history and understand contextual information in order to correctly deal with information needs, and finally the models need to be aligned with human values and have a certain level of empathy. Although the emergence and development of LLMs hold promise for interactive medical systems where patients can use these models to assess and solve personal health problems, there are certain potential risks and limitations due to the possibility of making incorrect judgments and decisions, and thus they cannot be directly applied to real-world scenarios, and the issue of safety in the field of medicine is even more important, and so interactive medical systems need to be further improved and developed [8]. Due to the importance of the medical field, scholars will continue to research the application of LLMs in the medical field.

(3) The application of LLMs in education. The clusters related to this research hotspot mainly contain artificial intelligence and bilingual education, and the cluster size is large. Artificial intelligence, as an important driving force leading the new round of technological revolution and industrial change, has given rise to a large number of new products, new technologies, new business models, and new modes. It has also brought more possibilities for the development of education, such

as teachers arranging e-homework through cloud platforms, using data to analyze the learning behaviors of students in the classroom, advancing the management process of the school towards digitization, and applying powerful arithmetic power to make the school management more accurate and optimized toward high-quality digital educational resources, promoting educational equity, etc. [18]. But, at the same time, scholars are also aware of the pitfalls of using LLMs in higher education. The improper use of AI can lead to academic plagiarism and other behaviors. However, in the current era of rapid technological development, maintaining a correct attitude towards AI and using AI reasonably can improve students' learning efficiency and ensure academic quality, so higher education institutions should encourage students to try to use LLMs such as ChatGPT creatively [18]. In addition, at the International Conference on Artificial Intelligence and Education held in 2022, government officials, experts, scholars, frontline teachers, and business representatives from dozens of countries around the world met to focus on the bright prospects of education development in the era of artificial intelligence, pointing out that it is of great significance to study the application of LLMs in education.

(4) Measurement and evaluation of LLMs. LLMs play an important role in research and daily use, and it is becoming increasingly important to evaluate and test them. Through the keyword clustering mapping, we can find that measurement invariance is related to this research theme, and "model", "quality", "aggregation operators", "2-tuple linguistic neutrosophic sets", and other co-occurring keywords are also closely related to the testing and evaluation of LLMs. Measurement invariance means that measurements obtained in different situations are similar or equivalent. Measuring the invariance of LLMs plays a crucial role in their own refinement and optimization [19]. In the field of artificial intelligence, testing the invariance of LLMs starts with the creation of a baseline model that does not take into account any measurement invariance constraints. For example, word co-occurrence patterns in a language corpus contain a large amount of conceptual knowledge, and in order to evaluate and improve the accuracy of a LLM in predicting contextual words in the corpus, researchers have continuously trained the model to improve its performance on a variety of semantic tasks so that the model was refined and thus achieved better prediction results [20]. Mannam et al. [21] also mentioned that models applied in the medical domain still need further research to optimize their performance for medical education and patient care. Similarly in the field of customer journey management, AI-based deep learning is needed to continuously optimize management models to provide valuable insights and practical guidance to decision makers [22]. Therefore, it is of great importance to test and evaluate LLMs.

Comprehensively analyzing the above research hot topics, we can see that the current core of scholars' attention to LLMs not only stays in the technology itself, but also focuses on its application in different aspects and practical value, that is how to improve the performance of LLMs for the happiness of human beings.

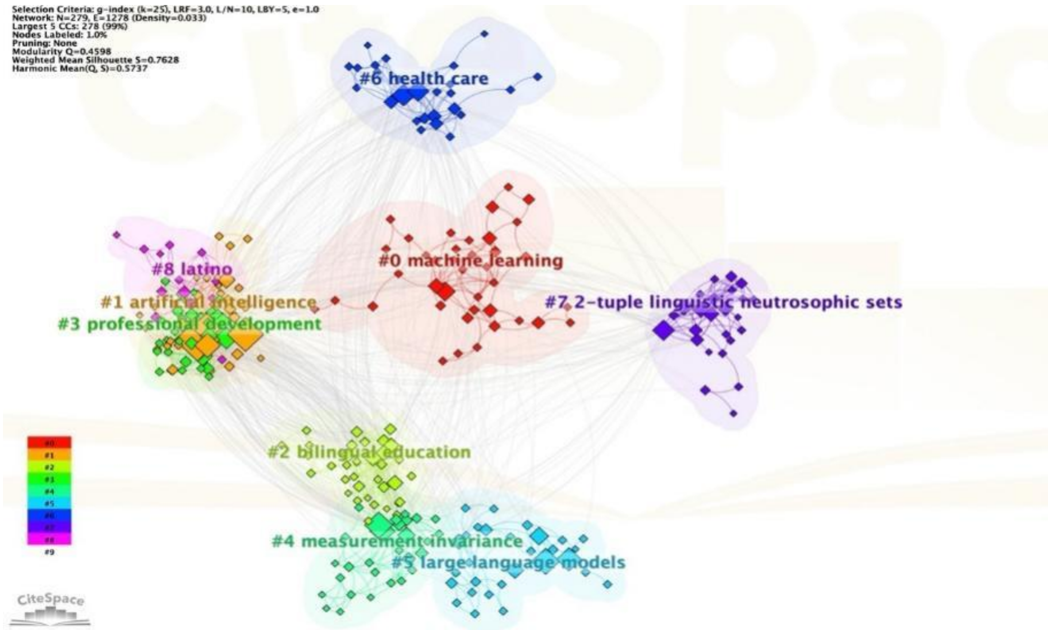

Figure 1. Keyword clustering mapping for LLMs.

Table 1. A clustering table of key words in LLMs.

Cluster ID	Label	Size	Silhouette	Mean (Year)	Top Terms
0	machine learning	78	0.734	2024	machine learning, automated analysis, health education, full-sentence defining style, artificial neural network, artificial intelligence, Large Language Model, stable diffusion
1	artificial intelligence	74	0.710	2021	education, dementia, childhood, attainment, cohort, second language, phonological similarity, speech perception, speech production
2	bilingual education	70	0.806	2022	bilingual education, two-way immersion, dual language, multilingual pedagogy, language awareness, second language, first language
3	professional development	64	0.754	2020	teacher education, professional development, language learners, 1-early childhood, university students, science education
4	measurement invariance	62	0.767	2023	struggling learners, theoretical perspective, phonemic awareness, comprehension instruction, language learners, model, validation, integration
5	LLMs	58	0.835	2021	artificial intelligence, image analysis, visual language model, language modal, young child feeding natural language processing
6	health care	52	0.689	2022	health care, impact, opportunity, science, systems, linguistic diversity, cross-sectional study, clinical practice, nursing student
7	2-tuple linguistics neutrosophic sets	49	0.880	2022	2-tuple linguistic neutrosophic sets, multiple attribute group decision, topsis method, political education quality
8	Latino	34	0.791	2022	Reading fluency, transparent orthography, reading accuracy, reading comprehension, linguistic comprehension, linguistic references

Existing research on LLMs presents the distinctive feature of being dominated by developed countries, while institutional collaboration is the main focus. Research related to the application of LLMs shows continuous growth. In order to identify scientific collaboration in the field of LLMs, this paper further uses Citespace to conduct collaborative network analysis[2]. The analysis results show that the United States has the largest contribution rate[3]. It has the largest number of publications, up to 882, accounting for more than 1/3 of the total sample of the literature analysis, and it can be said that the USA has a high scientific output capacity. China, Spain, the United Kingdom, Canada, Germany, Italy, Australia, and South Korea followed. The total number of publications from these eight countries exceeds 65% of the total literature analyzed in this study. This result is consistent with the centrality[4] values shown in Table 2. Centrality is a measure of the importance of a node's position in the network in terms of the number of shortest paths through the node. If the centrality value is higher than 0.1, the node can be considered as a critical node in the center of the network. Among the countries listed in Table 2, the centrality value of the United States is the highest, reaching 0.4. In addition, the centrality values of the United Kingdom, Spain, China, Canada, and Italy are all higher than 0.1, which indicates that the quality of the articles issued by these countries is high, and they are important nodes connecting the national cooperation in this research field.

At the same time, the related issues of the research are also the focuses of attention of international scholars. The United States particularly has the most fruitful achievements and influence in the research field of LLMs.

Table 2. Publication countries for LLMs.

Number	Count	Centrality	Year	Country
1	882	0.40	2020	USA
2	436	0.18	2020	China
3	245	0.22	2020	Spain
4	234	0.30	2020	England
5	156	0.10	2020	Canada
6	152	0.06	2020	Germany
7	140	0.10	2020	Italy
8	128	0.08	2020	Australia
9	97	0.06	2020	South Korea

Note: This table was analyzed by the authors using Citespace software to measure the dominant countries. The centrality refers to the intermediary role played by the country in scientific collaboration in this field; the number of publications mainly reflects the country's enthusiasm and attention to research in this field, and the year refers to the year of the country's first publication in the field of LLMs.

In terms of institutional and author cooperation, the cooperation between academic institutions is relatively close, and a certain range of cooperation network team has been formed in which mature

[2] Collaboration network analyses in Citespace can present the closeness of academic collaboration in the research area at different levels.

[3] The contribution of the dominant country in this study is measured according to the size of the centrality value, which is obtained by the Citespace software after analyzing the number of articles sent by the country, and if the value of the centrality is higher than 0.1, it can be regarded as a key node in the center of the network and the quality of the publication articles is high.

[4] The Citespace analysis of the literature data collected from the dominant countries allows us to obtain the centrality of different countries in the field of research, which is an important indicator of the level of importance and represents the pivotal role played in the field of research.

academic institutions are mostly dominated by major universities, including the University of California System, State University System of Florida, Cambridge, National University of Singapore, University of London, and so on. The cooperation between research authors is relatively loose, mainly in a small range of scientific collaboration. The number of network lines between academic authors is relatively small. And, there is no certain cooperation network group, which shows that scholars in the field of LLMs research are more inclined to cooperate with academic institutions to carry out research. They choose this method of collaboration to better realize the sharing of resources and improve the efficiency and quality of scientific research. Comprehensive analysis of the relevant data and maps confirms that the United States is in a leading position and plays an important intermediary role in the formation of the cooperative network, while other scholars would prefer to choose to cooperate through academic institutions.

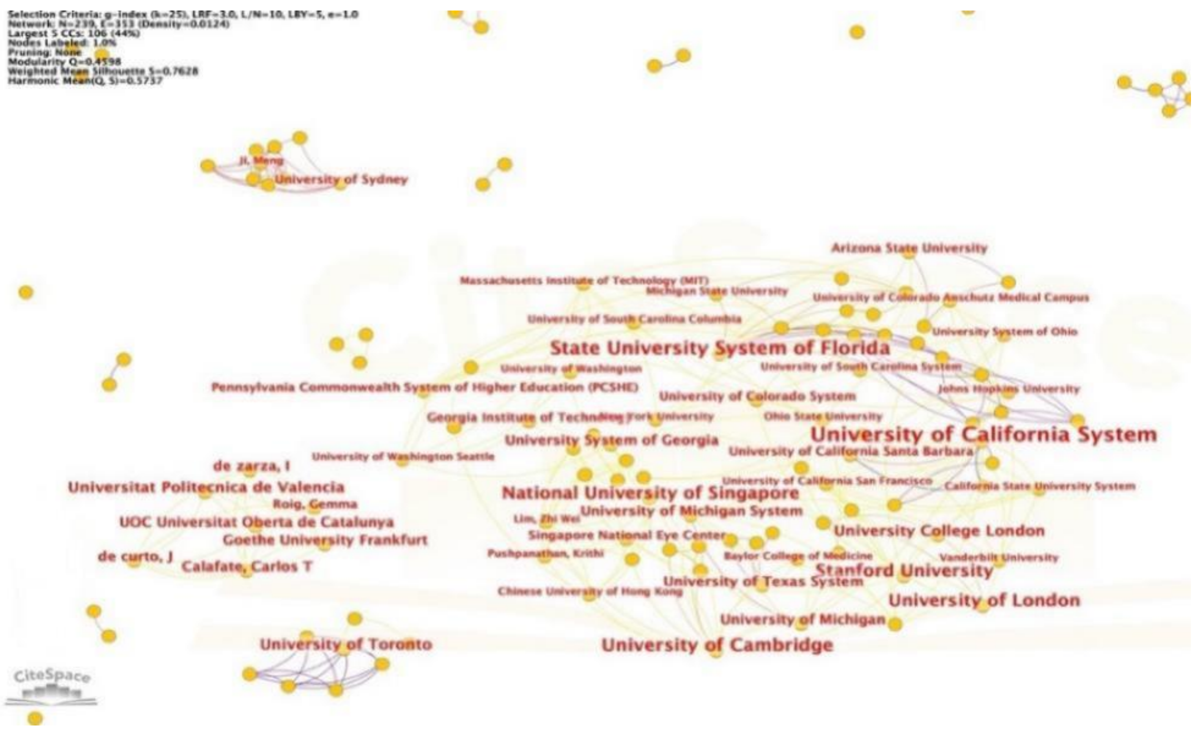

Figure 2. Collaborative network map of research institutions and research authors of LLMs.

Throughout the development of LLMs, it can be found that scholars have put more efforts into applied research, focusing on the research of application areas. Therefore, the study of the impact of scientific collaboration on the application of LLMs is in line with the current development trend, meets the needs of reality, and is conducive to scholars' understanding of the specific impact mechanism of scientific collaboration on the application of this field [23], so as to better promote the application of LLMs in different fields. Therefore, on the basis of previous studies, this paper will use a mixed linear regression model to assess the impact of scientific collaboration on the application of LLMs. Meanwhile, for some potential influencing factors, including financial support, dominant countries are also considered in the research and then the heterogeneous impact of scientific collaboration on the application of LLMs is deeply explored.

3. Models, variables, and data

Each paper is an independent observation that is not repeated in the sample. When confronted with a sample consisting of many independent and unduplicated papers, our focus naturally turns to the overall sample of which these papers are a part: the overall effect of scientific collaboration on the effectiveness of the application of the LLMs. We did not intend to delve into the specific effects of scientific collaboration in a single paper or a specific group of papers, but rather we expected to explore the general role of scientific collaboration on the effectiveness of the application of LLMs across the entire research field. To achieve this goal, we chose the Pooled Linear Regression Model (PLRM) as our analytical tool. The advantage of the PLRM is its ability to treat the data from all papers as a whole, assume that they all follow the same regression equation, and finally give an overall estimate of the effect of scientific collaboration on the application of LLMs.

In this study, a detailed description of the variables is given below:

Explained variable: Large Language Model application. LLMs is a natural language processing model based on deep learning technology, which has massive parameters and powerful text processing capabilities. LLMs usually learn the statistical laws and knowledge of language by pre-training on large-scale textual data so that they can generate natural and coherent text and understand reasoning about the text. The core of LLMs lie in their large-scale number of parameters, which enables the model to capture more complex linguistic phenomena and richer semantic information. Through the pre-training process, the LLM learns rich linguistic representations and contextual information, enabling it to handle a variety of complex natural language tasks, such as text generation, question and answer, text summarization, and sentiment analysis. We selected a series of keywords related[5] to LLMs around various aspects such as definition, methods or techniques, and application scenarios. After identifying these keywords, we further refined them by extracting the core words among them, calculating the number of occurrences of the core words in the literature, and using this as a characterization index for the application of LLMs.

Explanatory variable: scientific collaboration. This study uses two indicators, the number of authors and the number of institutions, to characterize this variable. The indicator of the number of authors directly reflects the number of authors involved in collaboration in a given research project. The number of authors usually reflects the breadth and depth of the collaboration. A larger number of authors may mean that the project covers a wider range of research areas and perspectives, and may also mean that more resources and experiences are integrated and utilized. In addition to the number of authors, the number of institutions is also an important indicator of scientific collaboration. It represents the number and diversity of research institutions involved in the collaboration. Different research institutions often have different research specialties and resource advantages, and through inter-institutional cooperation, resource sharing, and complementary advantages, mutual benefits can

[5] **Concept-related keywords**: Large Language Model, Language Understanding, Computational Linguistics, Transformer Model, Context-Aware Modeling, Cross-lingual Modeling, Multilingual Modeling, Pre-trained Language Model, Bidirectional Encoder Representations from Transformers (BERT).
Endemic technology-related keywords: Natural language processing (NLP), Long Text Processing, Context-aware, Language data, Generative Pre-trained Transformers (GPT), Corpus, Recursive Neural Network (RNN), Language Model Optimization, Semantic Analysis, Syntactic Analysis, Semantic Representation, Contextual word embeddings, Seq2Seq Models, Semantic Similarity, Named Entity Recognition (NER), Long-Document Handling, Masked Language Modeling, Language Model Perplexity.
Application scenario-related Keywords: Machine translation, Question answering, Chatbots, Content creation, Sentiment Analysis, Contextual Understanding, Knowledge Reasoning, Text Generation, Speech Recognition, Question Answering Systems, Language Generation, Text Classification.

be realized, which in turn promotes the application of LLMs.

Control variables. In this study, the number of article pages and year are chosen as control variables. The number of article pages is a measure of the length and level of detail of an article. Longer articles may contain more content and analysis, which may affect the application of LLMs. In addition, the technology and application of LLMs may continue to advance and evolve over time, and these technological advances may have a significant impact on the application of LLMs.

As a result, this paper constructs the following regression model:

$$Wordfreq_i = Naffiliation_i + Nauthor_i + Npage_i + Control_i + \varepsilon_i, \tag{1}$$

where $Wordfreq_i$ represents the number of applications of LLMs in paper i, $Naffiliation_i$ represents the number of institutions in paper i, $Nauthor_i$ represents the number of authors in paper i, $Npage_i$ represents the number of pages in paper i, $Control_i$ is the control variable of paper i, ε_i is the random fluctuation term.

In recent years, the number of published academic papers related to LLMs has shown a rapid growth trend. According to the detailed statistics of Google Scholar, we can see this remarkable change: Before 2020, the number of academic paper entries related to LLMs totaled about 17,800, which indicates that this field has some attention in the academic community. However, since 2020, the field has seen a dramatic rise in popularity, with a surge in the number of related academic paper entries to a whopping 74,200. This remarkable growth not only reflects the importance and attention of LLMs in academia, and, more importantly, this growth trend has continued to accelerate in subsequent years, demonstrating the prosperity and vibrancy of research in this field. In particular, in the first four months of 2024 (as of April 2024), the number of published papers has reached a whopping 31,200 despite the relatively short time span. This phenomenal growth rate suggests that the study of LLMs is in a phase of rapid development. In order to ensure the timeliness of this study and avoid the influence of outdated information, the papers related to LLMs during the period of January 2020 to April 2024 are selected as the research samples in this paper. The sample data come from two sub-collections: the Science Citation Index Expanded (SCI-EXPANDED) and Social Sciences Citation Index (SSCI), in the Web of Science database, with a total of 4610 data items. We excluded samples with missing values for key variables. At the same time, the data were truncated at 1% and 99% to avoid the effect of outliers and extreme values. The sample capacity after the above treatment is 2470.

In this paper, the sample is divided into five groups: full sample, developed and developing countries, and with financial support and without financial support, and then descriptive statistics are performed on the variables of different groups.

By looking at Table 3, we can find that the data distributions of these variables are characterized to varying degrees under different conditions. Based on the proportions of each sample group, the proportion of papers related to LLMs is significantly larger in developed countries than in developing countries, and the proportion of papers with financial support far exceeds the proportion of papers without financial support. At the mean level, the average number of LLM applications is about 42. Specifically, for the sample group of collaboration-led country types, we find that the mean level of the number of LLM applications in papers from developing countries is lower than that of developed countries. In the comparison of the sample groups divided into those with and without financial support, the number of LLM applications for papers with financial support is nevertheless lower than the number of applications for papers without financial support. Further, looking at the standard deviation, the overall sample has a larger standard deviation in the number of LLM applications, which indicates that there are significant differences in the number of LLM applications across papers. In particular, the level of variation in the number of LLM applications is greater for papers from developing countries and without financial support papers. In terms of the skewness coefficients, we find that the skewness coefficients of several variables are greater than 0, indicating that the data distribution exhibits some degree of right skewness, meaning that the number of applications is skewed towards the large-valued portion of the right-hand side. In particular, the skewness coefficients of Nauthor and Naffiliation are very large in the without financial support sample, which may be due to the fact that a small number of documents are unusually high in terms of the number of authors and the number of institutions, resulting in a right-skewed overall distribution. As for the kurtosis coefficients, the kurtosis coefficients of several variables are greater than 3 (the kurtosis coefficient of a normal distribution is 3), indicating that the data distribution has a spiky shape. In particular, the kurtosis coefficients of Wordfreq and Nauthor are very high in the full sample and in some subsamples, which may imply that there are extreme values in these variables, leading to a spiky distribution shape.

Table 3. Descriptive statistic.

	Variable	Mean	SD	Min	Max	Range	Skewness	Kurtosis
Full Sample (N=2494)	*Wordfreq*	42.070	61.770	0	570	570	2.626	12.710
	Nauthor	4.349	5.325	0	143	143	10.370	213.100
	Naffiliation	3.325	5.348	0	121	121	8.681	143.400
	Npage	16.490	9.557	1	217	216	6.095	99.170
Developed Countries (N=1506, P=60.385%)	*Wordfreq*	37.430	56.630	0	570	570	2.823	14.990
	Nauthor	4.214	3.929	0	36	36	2.748	14.880
	Naffiliation	3.410	4.262	0	59	59	3.334	26.920
	Npage	16.460	9.012	1	120	119	2.581	19.640
Developing Countries (N=988, P=39.615%)	*Wordfreq*	49.130	68.300	0	507	507	2.357	10.280
	Nauthor	4.555	6.929	0	143	143	11.070	184.200
	Naffiliation	3.195	6.672	0	121	121	10.000	143.400
	Npage	16.530	10.340	1	217	216	9.560	165.700
With Financial Support (N=1872, P=75.060%)	*Wordfreq*	39.260	57.960	0	477	477	2.378	10.280
	Nauthor	4.623	4.776	0	70	70	5.417	54.020
	Naffiliation	3.639	5.210	0	86	86	6.065	71.070
	Npage	16.840	9.878	1	217	216	7.021	113.700
Without Financial Support (N=622, P=24.940%)	*Wordfreq*	50.51	71.43	0	570	570	2.870	14.27
	Nauthor	3.524	6.647	0	143	143	15.54	315.8
	Naffiliation	2.381	5.644	0	121	121	15.37	315.7
	Npage	15.44	8.439	1	64	63	1.480	6.959

4. Empirical analysis

In this paper, we constructed a regression model based on Model 1 in the full sample state, and the parameter estimation method used is the ordinary least squares method. By analyzing Table 3, we find that there are differences in the performance of the sample data for the LLM application. In addition, the model setting and the selection of variables may not fully capture all the factors affecting the application of the LLM, leading to the existence of heteroskedasticity in the residuals of the regression model. The presence of heteroskedasticity not only affects the statistical inference of the regression model, but also reduces the estimation precision, making it difficult to accurately assess the impact of scientific collaboration on the application of LLMs. To cope with the problems posed by heteroskedasticity, this paper adopts the weighted least squares (WLS) method to estimate the model parameters. By giving different weights to different observations, WLS is able to reduce the effect of heteroskedasticity to a certain extent, thus providing more accurate parameter estimation and more reliable statistical inference. The parameter estimation results are shown in Table 4. The first and second columns show the parameter results obtained using ordinary least squares (OLS) estimation, where the second column shows the parameter estimates after controlling for time effects. The third and fourth columns, on the other hand, are the parameter results obtained using WLS estimation, where the fourth column is also the parameter estimates after controlling for time effects.

Table 4. Estimation results of the role of cooperative networks in the application of LLMs in full samples.

	OLS		WLS	
	(1)	(2)	(3)	(4)
Variables	Wordfreq	Wordfreq	Wordfreq	Wordfreq
Naffiliation	-0.468***	-0.446***	-0.467***	-0.450***
	(0.049)	(0.047)	(0.049)	(0.047)
Nauthor	0.041	0.004	0.040	0.004
	(0.061)	(0.060)	(0.061)	(0.060)
Npage	0.393***	0.456***	0.395***	0.449***
	(0.068)	(0.066)	(0.068)	(0.066)
Constant	2.070***	-0.138	2.063***	-0.117
	(0.200)	(1.522)	(0.199)	(1.560)
Time effects	No	Yes	No	Yes
Observations	2,494	2,494	2,494	2,494
Adjusted R^2	0.066	0.118	0.065	0.119

Notes: Standard errors in parentheses; ***, **, * represent $p<0.01$, $p<0.05$ and $p<0.1$, respectively.

Excessive institutional involvement limits the research and application of LLMs, and the effect of the number of authors on the application of LLMs is not significant. Both OLS and WLS estimation results show that the regression coefficients corresponding to the number of cooperating institutions range between -0.468 and -0.446, and both of them pass the significance test of 1%. This indicates that there is a significant negative relationship between the number of cooperating institutions and the application of LLMs. This may be due to the fact that too many collaboration institutions may lead to problems such as increased communication costs, fragmentation of research objectives, or intellectual property disputes, which may affect the efficiency of the development and application of LLMs [24]. On the other hand, the estimation results of OLS and WLS show that the regression coefficient corresponding to the number of authors is close to 0 and fails the significance test. This suggests that the number of authors is not a key factor in determining the number of applications in the research and application of LLMs. The possible reason is that the research and application of LLMs rely more on factors such as technical strength, algorithm optimization, and data quality, rather than mere numerical superiority [25]. Meanwhile, both OLS and WLS estimation results show that the regression coefficient of the number of dissertation pages is greater than 0, and both of them pass the 1% significance level test. This indicates that the number of dissertation pages shows a significant positive relationship with LLM applications. More detailed paper descriptions and more in-depth research analyses tend to increase the number of LLMs applications because they provide more information about model details, experimental methods, and theoretical foundations.

The application and development of LLMs may be affected by a variety of factors, such as scientific collaboration partners and financial support, which leads to the fact that scientific collaboration may show different effects in promoting the development of LLMs [26]. Collaboration with targets with a higher level of economic development can usually have a relatively well-developed technical infrastructure [27], such as high-performance computing centers and large-scale data centers, etc., which are advanced facilities that can provide strong material support for the research and development of LLMs. However, for academic partners with a lower level of economic development, their technological infrastructure is often relatively weak, and they lack high-performance computing equipment and large-scale data processing capabilities [28], which makes them often face many challenges in the development and application of LLMs. Financial support often plays an important role in technology application and innovation. Adequate financial support can ensure that technical teams have advanced computing equipment, large data sets, and specialized R&D staff to drive rapid technological advancement and innovation. Lacking the necessary financial support, they may not be able to purchase expensive computing equipment, collect and process large-scale data sets, and support long-term and in-depth research.

In order to examine the heterogeneous effects of publication countries (developing and developed countries) and financial support on the role of scientific collaboration in LLMs, this paper examines the impact of scientific collaboration on the application of LLMs after dividing the sample group based on the attributes of publication countries (based on the division of International Monetary Fund) and financial support. The results of the heterogeneity analysis are shown in Tables 5 and 6.

Table 5. Heterogeneity analysis based on financial support.

	Without funding		With funding	
	(1)	(2)	(3)	(4)
Variables	Wordfreq	Wordfreq	Wordfreq	Wordfreq
Naffiliation	-0.342***	-0.344***	-0.484***	-0.462***
	(0.098)	(0.096)	(0.056)	(0.055)
Nauthor	0.048	-0.004	0.072	0.032
	(0.107)	(0.105)	(0.075)	(0.073)
Npage	0.545***	0.652***	0.361***	0.391***
	(0.115)	(0.113)	(0.084)	(0.082)
Constant	1.779***	1.000***	2.054***	0.039
	(0.321)	(0.348)	(0.255)	(1.535)
Time effects	No	Yes	No	Yes
Observations	622	622	1,872	1,872
Adjusted R^2	0.055	0.106	0.064	0.113

Notes: Standard errors in parentheses; ***, **, * represent p<0.01, p<0.05 and p<0.1, respectively.

Table 6. Heterogeneity analysis based on publication countries.

	Developed		Developing	
	(1)	(2)	(3)	(4)
Variables	Wordfreq	Wordfreq	Wordfreq	Wordfreq
Naffiliation	-0.502***	-0.482***	-0.376***	-0.341***
	(0.060)	(0.058)	(0.083)	(0.080)
Nauthor	0.125*	0.104	-0.151	-0.212**
	(0.076)	(0.074)	(0.103)	(0.100)
Npage	0.291***	0.383***	0.605***	0.614***
	(0.082)	(0.080)	(0.123)	(0.119)
Constant	2.184***	0.062	1.794***	1.344***
	(0.247)	(1.490)	(0.344)	(0.350)
Time effects	No	Yes	No	Yes
Observations	1,506	1,506	988	988
R-squared	0.068	0.126	0.067	0.121

Notes: Standard errors in parentheses; ***, **, * represent $p<0.01$, $p<0.05$ and $p<0.1$, respectively.

With or without financial support does not change the role played by the collaborative network in the development of LLM applications. In Table 5, the coefficients on the number of institutions collaborating are all negative and all pass the 1% significance level test. This indicates that an increase in the number of institutions is negatively associated with the application of LLMs regardless of whether financial support is available. None of the coefficients for Nauthor pass the test of significance at a level of 1%, which indicates that an increase or decrease in the number of authors does not significantly affect the application of LLMs regardless of whether financial support is available. The coefficients for *Npage* are all positive and significant at the statistical test, which suggests that the number of pages, regardless of financial support, increases in the number of pages and is positively correlated with the adoption of LLMs, with or without financial support. This is consistent with common sense, as longer documents usually contain more vocabulary and more in-depth discussions.

By looking at the above results, it can be seen that there is no significant heterogeneity in the development of the application of LLMs with and without financial support. This may be related to the widespread impact of the digital economy [29]. The digital economy provides researchers with almost equal access to resources and information through its global network, efficient resource allocation, and extensive technology exchange [30]. As a result, even in the absence of direct financial support, researchers are able to leverage open source technologies, shared data, and online collaborative platforms of the digital economy to conduct high quality research and development. This trend towards equalization diminishes the decisive role of financial support in LLM research, making the results more dependent on the researcher's ability to innovate, the quality of the data, and the optimization of the algorithms than on the amount of financial support.

Differences in the publication countries lead to a slight heterogeneity in the role of cooperative networks on the application of the LLMs. Observing the regression results in Table 6, it can be found that the coefficients corresponding to *Naffiliation* are all negative and pass the significance test. *Nauthor* has a positive coefficient corresponding to the sample of developed countries and a negative coefficient corresponding to the sample of developing countries, and *Npage* has a positive coefficient corresponding to the sample of both groups and passes the significance test. This suggests that, in terms of the number of institutions, whether it is a cooperative network dominated by developed or developing countries, too many institutions will inhibit the application of the LLMs to a certain extent. The differences are mainly in terms of the number of collaborators. In developing country-led scientific collaboration, an excessive increase in the number of collaborators inhibits the application of LLMs; meanwhile, in developed countries, there is a less significant positive correlation between the increasement in the number of collaborators and the application of LLMs, meaning that the increase in the number of collaborators has a certain effect on the application, but this effect is not very obvious. Developing countries are relatively more limited in terms of research resources, technological infrastructure, and specialized personnel. When the number of collaborators increases rapidly, the limited resources may be dispersed excessively, resulting in fewer resources available to each collaborator, which makes it difficult to effectively promote the research, development, and application of LLMs [31]. In addition, communication and coordination among collaborators may become more complex and difficult, increasing the friction and cost of collaboration, and further affecting the effectiveness of the application of LLMs [32]. Developed countries have relatively well-developed resources and technology bases, and increasing the number of collaborators can facilitate the application of LLMs to a certain extent. This may be due to the fact that more collaborators can bring in a wider range of resources, technology, and experience, which enhances the strength and innovativeness of the research and development of LLMs [33]. However, this positive effect is not quite as pronounced, possibly because developed countries already have a more mature research system and technology base, and the marginal benefits of increasing the number of collaborators is relatively small [34].

5. Discussion and conclusions

This paper briefly introduces the application of LLMs in various fields, and then takes the education field as an example and lists a series of relevant policy and documents that issued to promote the development of LLMs in the education field. Then, this paper describes the development of research areas, research hotspots, and research characteristics of LLMs through the literature analysis software Citespace. This paper then empirically examines the relationship between scientific collaboration and the role of LLMs. Considering that both financial support and the country of publication can affect the relationship, this paper further analyzes the heterogeneity. The empirical results show the following: First, the increasing number of authors does not have a significant impact on the application of LLMs, and even when too many institutions are involved in the cooperation, it may have a certain limitation on the application of LLMs. Second, the factor of resource support does not have a heterogeneous impact on the application of the development of LLMs. Lastly, the dominant country of the issuance of the article will bring a degree of heterogeneity on the impact of the application of LLMs, but the impact is relatively mild.

Based on the previous analysis, this paper puts forward the following suggestions to promote the application and development of LLMs. First, the structure and function of scientific collaboration institutions should be optimized. In order to ensure a balance between the quality and quantity of participating institutions, a strict access mechanism should be set up for academic institutions participating in LLM research. The mechanism should cover multiple dimensions such as scientific research strength, research experience, resource investment, and so on. At the same time, a small number of core teams should be encouraged to lead the research and application of LLMs, while other institutions should act as collaborators so as to avoid the dispersion of resources and the reduction of efficiency caused by the participation of too many institutions. Second, the efficiency and flexibility of the use of financial support should be improved. When providing financial support, the purpose and scope of the use of financial support should be clarified to ensure that the financial support can be used to promote scientific collaboration and applied research on LLMs. In addition, scholars should be given a certain degree of autonomy in the use of financial support, allowing them flexibly to adjust how they use the financial support according to the actual needs of their research projects, so as to improve the efficiency of the use of financial support. Third, international scientific and technological collaboration should be strengthened. Scientific collaboration should be actively sought with countries with strong technical strength, large dataset sizes, and experienced industry applications to introduce advanced technical concepts and solutions, obtain rich training data, and accelerate the application of LLMs in specific fields. At the same time, attention should be paid to cultivating scientific research talents with cross-cultural communication skills and enhancing their communication and coordination abilities in international cooperation, so as to promote international exchange and cooperation in LLM technology.

Author contributions

Suyan Tan: Writing – review & editing, Supervision, Conceptualization. **Yilin Guo**: Writing – original draft, Formal analysis.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was funded by the National Social Science Fund Project "Research on Cognitive Load in User Experience of College English Cloud Platform" (22BYY094), the Guangdong Postgraduate Education Innovation Project "Research on Innovative Mode of Postgraduate Moral Education Empowered by Information Technology" (2023JGXM_098), the Project of Philosophy and Social Sciences Planning of Guangdong Province (GD23WZXC01-13).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. J. Liu, D. Shen, Y. Zhang, W. B. Dolan, L. Carin, W. Chen, What makes good in-context examples for GPT-3? *Proceedings of Deep Learning Inside Out (DeeLIO 2022): The 3rd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures*, **3** (2022), 100–114. https://doi.org/10.18653/v1/2022.deelio-1.10

2. S. Karan, S. Azizi, T. Tu, S. S. Mahdavi, J. Wei, H. W. Chung, et al., Large Language Model encode clinical knowledge, *Nature*, **620** (2023), 172–180. https://doi.org/10.1038/s41586-023-06291-2
3. A. Laios, G. Theophilou, J. D. De, E. Kalampokis, The future of AI in ovarian cancer research: the Large Language Model perspective, *Cancer Control*, **30** (2023), 10732748231197915. https://doi.org/10.1177/10732748231197915
4. S. Ali, P. Chourasia, M. Patterson, When protein structure embedding meets Large Language Model, *Genes*, **15** (2024), 25. https://doi.org/10.3390/genes15010025
5. E. Dotan, G. Jaschek, T. Pupko, Y. Belinkov, Effect of tokenization on transformers for biological sequences, *Bioinformatics*, **40** (2024), 196. https://doi.org/10.1093/bioinformatics/btae196
6. G. Leng, Challenge, integration, and change: ChatGPT and future anatomical education, *Med. Educ. Online*, **1** (2024), 2304973. https://doi.org/10.1080/10872981.2024.2304973
7. M. L. Tsai, C. W. Ong, C. L. Chen, Exploring the use of Large Language Model (LLMs) in chemical engineering education: building core course problem models with Chat-GPT, *Educ. Chem. Eng.*, **3** (2023), 71–95. https://doi.org/10.1016/j.ece.2023.05.001
8. S. Griewing, N. Gremke, U. Wagner, M. Lingenfelder, S. Kuhn, J. Boekhoff, Challenging ChatGPT3.5 in senology–An assessment of concordance with breast cancer tumor board decision making, *J. Pers. Med.*, **10** (2023), 1502. https://doi.org/10.3390/jpm13101502
9. D. Schillinger, R. Balyan, S. A. Crossley, D. S. McNamara, J. Y. Liu, A. J. Karter, Employing computational linguistics techniques to identify limited patient health literacy: findings from the ECLIPPSE study, *Health Serv. Res.*, **1** (2020), 132–144. https://doi.org/10.1111/1475-6773.13560
10. I. R. Indran, P. Paramanathan, N. Gupta, N. Mustafa, Twelve tips to leverage AI for efficient and effective medical question generation: a guide for educators using Chat GPT, *Med. Teach.*, **2023** (2023), 2294703. https://doi.org/10.1080/0142159X.2023.2294703
11. C. B. Divito, B. M. Katchikian, J. E. Gruenwald, J. M. Burgoon, The tools of the future are the challenges of today: the use of ChatGPT in problem-based learning medical education, *Med. Teach.*, **3** (2024), 320–322. https://doi.org/10.1080/0142159X.2023.2290997
12. Y. Wang, K. Yan, Machine learning-based quantitative trading strategies across different time intervals in the American market, *Quant. Financ. Econ.*, **4** (2023), 569–594. https://doi.org/10.3934/QFE.2023028
13. K. Amin, P. Khosla, R. Doshi, S. Chheang, H. P. Forman, Artificial intelligence to improve patient understanding of radiology reports, *Yale J. Biol. Med.*, **3** (2023), 407–414. https://doi.org/10.59249/NKOY5498
14. S. Luo, W. Lei, P. Hou, Impact of artificial intelligence technology innovation on total factor productivity: an empirical study based on provincial panel data in China, *Natl. Account. Rev.*, **2** (2024), 172–194. https://doi.org/10.3934/NAR.2024008
15. A. Guleria, K. Krishan, V. Sharma, T. Kanchan, ChatGPT: ethical concerns and challenges in academics and research, *J. Infect. Dev. Ctries*, **9** (2023), 1292–1299. https://doi.org/10.3855/jidc.18738
16. S. Mestiri, Credit scoring using machine learning and deep learning-based models, *Data Sci. Financ. Econ.*, **2** (2024), 236–248. https://doi.org/10.3934/DSFE.2024009

17. P. Maddigan, T. Susnjak, Chat2VIS: generating data visualizations via natural language using ChatGPT, Codex and GPT-3 Large Language Model, *IEEE Access*, **11** (2023), 45181–45193. https://doi.org/10.1109/ACCESS.2023.3274199
18. S. Milano, J. A. McGrane, S. Leonelli, Large Language Model challenge the future of higher education, *Nat. Mach. Intell.*, **4** (2023), 333–334. https://doi.org/10.1038/s42256-023-00644-2
19. P. Theodorou, P. Theodorou, Valuation of big data analytics quality and competitive advantage with strategic alignment model: from Greek philosophy to contemporary conceptualization, *Data Sci. Financ. Econ.*, **1** (2024), 53–64. https://doi.org/10.3934/DSFE.2024002
20. C. Kauf, A. A. Ivanova, G. Rambelli, E. Chersoni, J. S. She, Z. Chowdhury, et al., Event knowledge in Large Language Model: the gap between the impossible and the unlikely, *Cognit. Sci.*, **47** (2023), e13386. https://doi.org/10.1111/cogs.13386
21. S. S. Mannam, R. Subtirelu, D. Chauhan, H. S. Ahmad, I. M. Matache, K. Bryan, et al., Large Language Model- based neurosurgical evaluation matrix: a novel scoring criteria to assess the efficacy of ChatGPT as an educational tool for neurosurgery board preparation, *World Neurosurg.*, **180** (2023), E765–E773. https://doi.org/10.1016/j.wneu.2023.10.043
22. T. C. Wang, Deep learning-based prediction and revenue optimization for online platform user journeys, *Quant. Financ. Econ.*, **1** (2024), 1–28. https://doi.org/10.3934/QFE.2024001
23. S. A. Gyamerah, C. Asare, A critical review of the impact of uncertainties on green bonds, *Green Financ.*, **1** (2024), 78–91. https://doi.org/10.3934/GF.2024004
24. H. Shin, K. Kim, D. F. Kogler, Scientific collaboration, research funding, and novelty in scientific knowledge, *Plos One*, **17** (2022), e0271678. https://doi.org/10.1371/journal.pone.0271678
25. O. N. E. Kjell, K. Kjell, H. A. Schwartz, Beyond rating scales: with targeted evaluation, Large Language Model are poised for psychological assessment, *Psychiatry Res.*, **333** (2024), 115667. https://doi.org/10.1016/j.psychres.2023.115667
26. Z. Williams, H. Apollonio, The causation dilemma in ESG research, *Green Financ.*, **2** (2024), 265–286. https://doi.org/10.3934/GF.2024011
27. Y. Wen, Y. Xu, Statistical monitoring of economic growth momentum transformation: empirical study of Chinese provinces, *AIMS Math.*, **10** (2023), 24825–224847. https://doi.org/10.3934/math.20231266
28. Z. Li, Q. Lai, J. He, Does digital technology enhance the global value chain position? *Borsa Istanb. Rev.*, 2024. https://doi.org/10.1016/j.bir.2024.04.016
29. Y. Gao, H. Yang, The measurement of financial support for real estate and house price bubbles and their dynamic relationship: an empirical study based on 31 major cities in China, *Natl. Account. Rev.*, **2** (2024), 195–219. https://doi.org/10.3934/NAR.2024009
30. Z. Li, B. Chen, S. Lu, G. Liao, The impact of financial institutions' cross-shareholdings on risk-taking, *Int. Rev. Econ. Financ.*, **92** (2024), 1526–1544. https://doi.org/10.1016/j.iref.2024.02.080
31. Z. Kohus, M. Demeter, G. P. Szigeti, L. Kun, E. Lukács, K. Czakó, The influence of international collaboration on the scientific impact in V4 countries, *Publications*, **10** (2022), 35. https://doi.org/10.3390/publications10040035
32. R. Cai, W. Tian, R. Luo, Z. Hu, The generation mechanism of research leadership in international collaboration based on GERGM: a case from the field of artificial intelligence, *Scientometrics*, 2024. https://doi.org/10.1007/s11192-024-04974-9

33. M. M. Danquah, O. B. Onyancha, B. K. Avuglah, Patterns and trends of university-industry research collaboration in Ghana between 2011 and 2020, *Inf. Discov. Deliv.*, 2024. https://doi.org/10.1108/IDD-11-2022-0122
34. B. B. Xu, S. Y. Liu, J. F. Guo, Graph-based algorithm for exploring collaboration mechanisms and hidden patterns among top scholars, *Expert Syst. Appl.*, **249** (2024), 123810. https://doi.org/10.1016/j.eswa.2024.123810