Andreea PERNICI, PhD Candidate (corresponding author)
andreea.pernici@csie.ase.ro
Bucharest University of Economic Studies, Bucharest, Romania

Stelian STANCU, PhD
stelian.stancu@csie.ase.ro
Bucharest University of Economic Studies, Bucharest, Romania
Centre for Industrial and Services Economics, Romanian Academy, Bucharest, Romania

Daniela-Elena MARINESCU, PhD
daniela.marinescu@csie.ase.ro
Bucharest University of Economic Studies, Bucharest, Romania

A Half-Century of Artificial Intelligence: A Global Mapping of Academic Research Trends and Leading Hubs

Abstract. *Artificial intelligence (AI) is without a doubt the term of the moment, a true revolution that should be understood not only from the technical perspective but from a geographical and evolutive one as well. Considering its clear momentum, the current paper aims to construct an analysis framework that can explore the AI academic path, in a 50-year window of time, integrating the past, current, and future perspectives. This will be done through a bibliometric analysis, having as a data source the entire collection of Web of Science AI publications. Furthermore, the paper will be split into three research objectives, starting with an initial exploration and prediction of the AI publication trend, a mapping of the global academic centres, and an in-depth exploration of the two leading hubs: the USA and China. Thus, using Python and VOSViewer, the current approach can distinctively underscore the dynamic nature of AI from its beginnings up to a very recent moment (2024), while also highlighting the domain breakdown, topic networks, and country specifics.*

Keywords: *artificial intelligence, bibliometric analysis, academia, Web of Science, USA, China, Python, VOSViewer.*

JEL Classification: C88, C22, I23, O33.

Received: 16 February 2025	Revised: 24 May 2025	Accepted: 2 June 2025

1. Introduction

Although the *Artificial Intelligence* (AI) concept traces back to ancient times, one of the pivotal moments for the academic narrative was the *Dartmouth Conference* in 1956, considered by many the birth of AI as a field of study. At the time, *John McCarthy*, along with several other computer scientists, attended an academic workshop in which they explored the idea that human intelligence could one day be replicated by machines. This idea turned them into the pioneers of AI, a newly designed field, which would become a true revolution 70 years later.

DOI: 10.24818/18423264/59.2.25.05

Therefore, especially in the last 3 years, AI has seen exponential growth in terms of visibility, awareness, investments, and, of course, research. In fact, the last element will be the main rationale behind this paper, namely extracting meaningful insights regarding the path that the AI academic interest has followed and evolved into, from a geographical and topical point of view. However, since it is difficult to cover in one article the immense volume of publications on the AI subject, we will limit our approach to the *Web of Science* (Clarivate) platform, first introduced in 1964. Eleven years have passed from that initial moment until the first publications classified on the topic of artificial intelligence, with the yearly volumes increasing exponentially ever since.

Having this in mind, we can now develop further our research goal: tracing the development of AI over the last 50 years and providing insights into its geographical distribution, the leading academic hubs, and domain specificities. Our paper will start with a trend exploration and prediction; afterwards, the attention will be switched to the more recent years, considering the proliferation of the subject. Without a doubt, the unfolding of the domain was due to the advancements that were pushed forward by private companies and investors, with major breakthroughs both in terms of actual hardware technology, such as the *graphics processing units* (GPU), and in terms of newly-designed and deployed algorithms: *neural networks* (NN), *deep learning*, *natural language processing* (NLP), or *big data,* being just some of the examples.

Considering this, in the second part of our paper we aim to investigate the diversity of approaches, from theoretical frameworks to practical applications, by employing a comprehensive bibliometric analysis. Since AI is already on the agenda of any stakeholder, either political, private, or scholarly, this framework could prove to be valuable in detecting its poles of power, their specificities, and potential interconnections.

2. Literature review

However, before proceeding with the exploration of the current model, we need to first assess the status quo of the related literature. Although we will limit our bibliography to similar bibliometric methodologies, we need to also mention some fundamental references for the history of AI, considering the evolutionary perspective of our approach.

Thus, as stated in the introduction, the AI origin is marked by the *Dartmouth Conference Proposal* (McCarthy et al., 1956), in which the artificial intelligence term was used for the first time. Going forward, we cannot discuss the beginnings of AI, without mentioning the name of Alan Turing, a clear symbol of the groundbreaking technical advancements that were to change the world as we know it. Through the *Turing Machine*, a theoretical construct that formalised the principles of computation and algorithmic processes, the mathematician set the foundation for the entire computer science dimension, defining the limits of what can be computed and *thought* by a machine. Moreover, in *Computing Machinery and Intelligence*

(Turing, 1950), the author introduces the idea of a *Turing Test*, a fundamental concept for AI, in which the objective is to evaluate a machine's ability to actually demonstrate human-like intelligence.

Going forward, crucial steps were also taken by Newell & Simon (1972), the authors exploring further several fundamental ideas of AI, such as *general problem solver* (GPS), the *logic theorist*, or *information processing language* (IPL), and applying it to cognitive psychology. From a more technical point of view, we need to mention Pearl (1988) who introduced *Bayesian networks*, Hinton, Rumelhart & William (1986) who redefined *neural networks* and *backpropagation* algorithms, or LeCun et al. (1988) who significantly contributed to the *convolutional neural network* (CNN) development, a cornerstone for deep learning and image processing.

Without a doubt, many others have impacted the field throughout the last 60 years, however, we will now focus on the more recent references which also showcase a similar methodology to the one exposed in this paper. Therefore, in our literature exploration, we managed to find a multitude of publications that either revolve around the subject of AI history or employ a bibliometric analysis in the process. Most of these references focus on a specific sub-domain, such as urban models (Maisonobe, 2022), chatbots (Liu & Duffy, 2023), supply chain (Riahi et al., 2021), business (Dissanayake et al., 2024) engineering applications (Shukla et al., 2019), education (Guan, Mou & Jiang, 2020) or physical, natural, life, and social sciences (Hajkowicz et al., 2023). At the same time, a consistent category of papers focuses on medical domains, such as dentistry (Xie et al., 2024), vascular surgery (Javidan et al., 2022), or the general evolution in healthcare and medicine (Tran et al., 2019).

Nevertheless, when we switched our attention to more generic papers, we managed to find very similar approaches. For example, Ho & Wang (2020) have designed a methodology focused on the Web of Science publications that included *artificial intelligence* as a topic in the title, abstract, and author keywords. Starting from that, the authors analysed the data from multiple perspectives, such as time evolution, research areas and categories, countries and institutions, high impact and citations, or general topics. In terms of results, the main insights will coincide with the ones generated in our endeavour, namely that the USA and China are the dominant contributors (along with Iran as a recent hub), while *neural networks*, *optimisation*, *support vector machines*, or *big data* have been the main technologies explored and preferred by scholars throughout the recent years (up until 2018).

Similarly, a paper undertaken by Gao et al. (2021) included a selection of 12,301 articles, that mark the top 10% of yearly high-citation ones in the period 2008-2018. The aspects analysed are once again comparative, with the authors presenting aspects such as the distribution of articles by year and country, collaborative networks between countries, high-frequency affiliations, and most importantly, topic conclusions using visualisation networks. At the same time, the article included a cluster analysis, as part of the *SAS software*, generating a total of 10 clusters of topics, each with their afferent keywords. Finally, a similar approach has also been described by Gao & Ding (2022), this time extending the universe of time to 20 years.

Therefore, although the actual paper does not propose a totally new approach, it can be seen as an extension of the current status quo, considering both the vast time frame selected and the focus on the 2024 year later on. Last but not least, it is important to mention that considering the limitations of space, we will not reference all the papers that are going to be used in the applicative part, however, the filters added to the Web of Science platform will be clearly identified.

3. Model specification

3.1 Model Description and Objectives

Reaching now the methodological part of our paper, the current framework relies on bibliometric analysis, a method that has been increasing in usage over the last years, due to its multi-valence functions. Statistics, mathematics, clustering, quantitative visualisations, graphs, and networks are just some of the instruments that can contribute to a valuable bibliometric analysis, especially in the context of AI research. Therefore, several of these will also be computed in the current paper, as enablers to our research goals.

Finally, in order for the information to be clearly structured and presented, we will now illustrate our three research objectives, each of them having a specific methodology behind it.

- **Objective 1:** Illustrate the AI-topics publication trend in a 50-year time frame and predict the volume of research output over the next 3 years.
- **Objective 2:** Map the geographical distribution of AI-topics research at a global level and highlight the leading scholar hubs.
- **Objective 3:** Deep-dive into the specifics of the leading hubs, and explore the differences in the AI publication subjects and clusters.

3.2 Methodology

Starting with the first objective, we will rely mainly on *Python* visualisation tools, considering as a data source the *Web of Science* (WoS) list of publications with AI as a topic, throughout the 1975-2024 period.

Next, for the prediction sub-goal, we will employ an *ARIMA* (Autoregressive Integrated Moving Average) model, one of the most used methods for time-series forecasting. Nevertheless, as a mandatory initial step, we need to first ensure the stationarity of our time series, and for this, we will compute both a *log transformation* (1) and a *first-differencing* procedure (2). Their goals will be to first stabilise the variance and transform the series into a normally distributed one, and then to generate a stationary dataset. At the same time, we need to be careful that after the ARIMA model has been built, the transformations require reverting, which will be done using the *exponential function* (1) and the *cumulative sum* (2).

$$Y_t' = log(Y_t) \quad reversed\ with => \quad Y_t = e^{Y_t'} \tag{1}$$

$$\Delta Y_t = Y_t - Y_{t-1} \quad reversed\ with => \quad Y_t = Y_{t-1} + \Delta Y_t \tag{2}$$

Lastly, in order to make sure that the data series is stationary, we have employed the *Augmented Dickey-Fuller (ADF)* test, which follows the below hypothesis.

- **Null Hypothesis** ($\boldsymbol{H_0}$): The time series is non-stationary (has a unit root).
- **Alternative Hypothesis** ($\boldsymbol{H_1}$): The time series is stationary (it does not have a unit root).

After ensuring the stationarity, we can finally proceed to building the ARIMA model, considering the (3) equation. We can therefore note Y_t, the time series at the moment t, the constant c, the autoregressive (AR) coefficients ϕ_i, the moving average (MA) coefficients θ_i, and ϵ_t, the white noise error. Finally, using this model, we will generate a three-year prediction for the overall volume of AI publications on the WoS platform, an important point of reference for its future trend.

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \tag{3}$$

Proceeding to the second objective, namely the geographical mapping, we will once again use *Python* visualisation packages, more specifically *geopandas*, while as a timeframe, we will not make any exclusions, since we aim to see the total contributions over the last 50 years.

Finally, the third objective will be the most complex, integrating multiple views, useful for an efficient appraisal of the country specifics. We will continue using *Python* visualisation tools, however, we will also map out the *co-occurrence* networks using the *VOSViewer* application. The tool has been intensively used in the related literature, due to its potential to design visualisations that can be highly parametrised based on the research needs.

To further describe it, the main principle behind a co-occurrence map is that it will calculate the distances between the pairs of words, and then codify them into a two-dimensional space, placing the more similar items closer together. In this process, each word (label) will be given an x and a y, or the *2D coordinates* of the visualisation, the method computing in the background the number of general occurrences, the links (or direct connections) with other items, and the total link strength. A more extensive definition of these terms is available in Table 1.

Table 1. VOSViewer indicators and definitions

Elements	Definition
Label	The actual name of the item (e.g., "Artificial Intelligence", "Deep Learning")
Occurrences	The number of times an item (label) appears in the dataset
x, y	The 2D coordinates of the item (label) in the visualisation
Links	The number of direct connections an item has to other items in the network
Total link strength	The sum of link strengths an item has with all other items (higher values mean stronger connections)

Source: Authors' own processing, based on VOSViewer indicators.

Finally, *VOSViewer* also performs a clustering technique, based on the similarity (distances) between the pairs of words. It uses *a Density-Based* algorithm in order to detect high-density areas in the co-occurrence mappings, in an automatically adaptive way. To put it differently, the application can generate an optimal number of clusters without being known a priori. This approach will help us in interpreting and characterising the leading hubs, based on the specificities that their research implies.

Therefore, considering this methodological flow, we can now proceed to the results section of our paper, with an individual assessment of each objective.

4. Results and discussion

Objective 1. Publication trend

Starting with the first objective, namely the illustration of the AI publication trend in the last 50 years, we have filtered all the WoS publications that have the term *artificial intelligence* as a topic, meaning that it is integrated into the title, abstract, keyword plus, or author keywords. Although this is an extensive approach, we have considered it a good starting point in understanding the true magnitude of the field, the query generating almost 220,000 results for the entire 1975-2024 period, coming from over 200 fields of study and 200 countries.

Next, in Figure 1 we have illustrated the publication trend, with a conclusive growth over the years. In a deep-dive approach, we can observe a very small volume of papers being published up until the 2010 decade, increasing swiftly afterward, and then surging completely near the 2020s. To get a glimpse of the numeric growth, the number of publications almost doubled from 2017 (4,000) to 2018 (7,000), and then again in 2019 (12,000). In addition, throughout the most recent years, the rhythm continued to increase, reaching an absolute pinnacle in 2024, with more than 50,000 papers (a quarter of the total) being published on the well-known platform. In comparison, in 2023 the number was set at approximately 37,600, while in 2022 at 31,700.

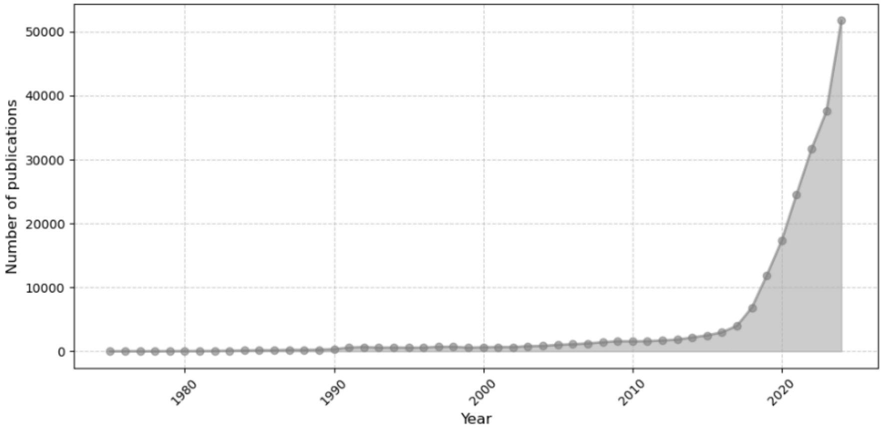

Figure 1. Evolution of WoS AI publication volumes, 1975-2024

Source: Authors' own processing.

Prediction of WoS publication volume

Now that we have gained an understanding of the evolution trend, the second part of our first objective is to predict the publication volumes for the next three years, up to 2027. We have therefore computed an ARIMA model that includes all the 49 previous data moments, applying both a log transformation and a first-differencing, until arriving at a negative *ADF statistic* (-3.915, confirmed by the 0.0019 *p-value*) and a stationary dataset.

Finally, the predicted numbers can be observed in Table 2, along with the actual values for the last three years. What we can notice is that there are some differences in the predicted values versus the actual ones, with the method forecasting a much smoother ascending trend between the three years, while in reality, the jump from 2022 to 2024 volumes was much more abrupt. However, if we were to look at the final 2024 status, the numbers are lining up, with a surplus in the actual trend. That could be easily explained by the explosion that AI has shown throughout the last year, with the model not being able to capture it in its entirety.

Table 2. ARIMA prediction of WoS AI publication volumes

Year	Actual Data	Prediction
2022	**31,681**	32,498
2023	**37,592**	41,128
2024	**52,908**	49,502
2025		67,425
2026		84,810
2027		**104,376**

Source: Authors' processing.

Going further into the future, we can see that if the trend persists, the number of papers published on the topic of AI will double versus 2024, reaching a forecast value of over 100,000 papers in 2027. Graphically, this conclusion is plotted in Figure 2.

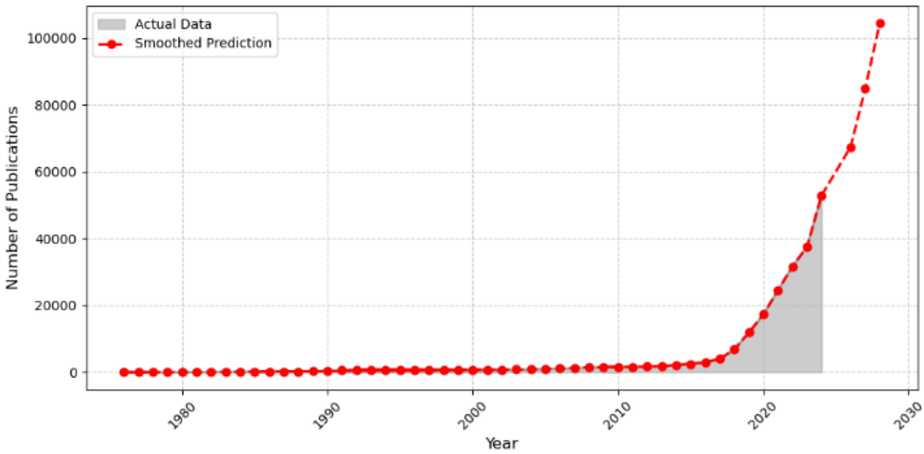

Figure 2. ARIMA prediction of WoS AI publication volumes, 1975-2027
Source: Authors' own processing.

To sum up, while the current exponential growth clearly highlights AI's momentum, we should approach this prediction with caution. Although technology is rapidly advancing and generating vast amounts of research, history suggests that all innovations eventually reach a plateau. Therefore, it remains to be seen whether AI will continue expanding at this fast pace or if it will follow the natural trajectory of technological progress.

Objective 2. A global mapping – geographical research hubs

After understanding, at least broadly speaking, the general evolution of the AI research volumes, we can now proceed to the second objective stated in the methodology: the exploration of the geographical distribution of academic research, in a mapping format. Thus, in Figure 3, we have plotted the number of WoS publications generated by each country in the entire time frame selected - 1975-2024.

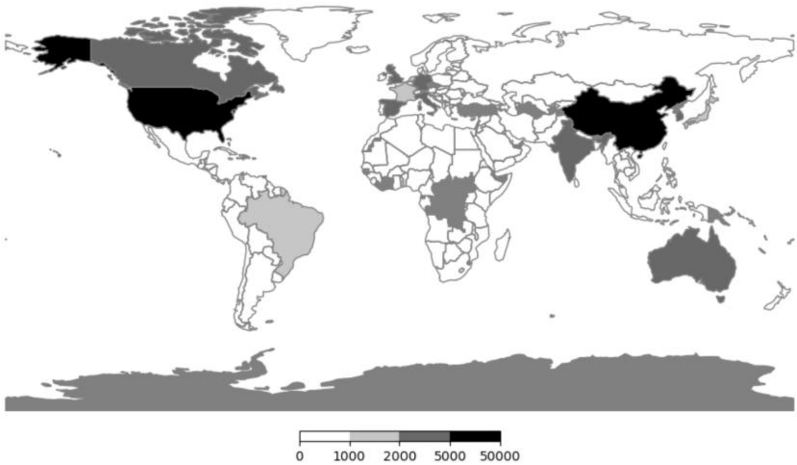

Figure 3. Geographical mapping – WoS AI publication volumes, 1975-2024
Source: Authors' own processing.

Without a doubt, the first thing we can notice is the prevalence of two global research hubs, the United States and China, for which the number of publications with artificial intelligence as a topic exceeds the 40,000 threshold. Thus, the two countries can be defined as the primary contributors to the research output, with very close total volumes throughout the entire period of time.

Besides those, however, we can also pinpoint some other important poles of interest, such as India, the United Kingdom, Germany, and Italy, all of these accounting for more than 10,000 papers on different AI topics over the last 50 years. The next cluster of performers will be those who exceeded the 5,000 publications, in this category, including Spain, Canada, South Korea, Australia, France, Japan, and Saudi Arabia. Thus, we can see a relatively diverse presence, with contributions spanning across multiple continents. At the same time, we need to also assess the direct correlation with the most economically potent countries, as well as the ones that have designed and implemented an efficient educational system, a factor that played a definitive role in the published outputs.

Objective 3. A comparative deep-dive into the leading hubs' characteristics

Finally, now that we have assessed the two major poles of interest when it comes to artificial intelligence academia, we can go ahead and deep-dive into both their similarities and differences. As a first step, we aim to assess the subdomains that the AI publications are in correlation to, trying to decipher whether there are distinctive patterns or priorities at a national level.

However, as an initial endeavour, it is important to also assess the categories from a global perspective, using the Web of Science classification. Therefore, in Figure 4 we have illustrated the most common categories, for the entire time frame.

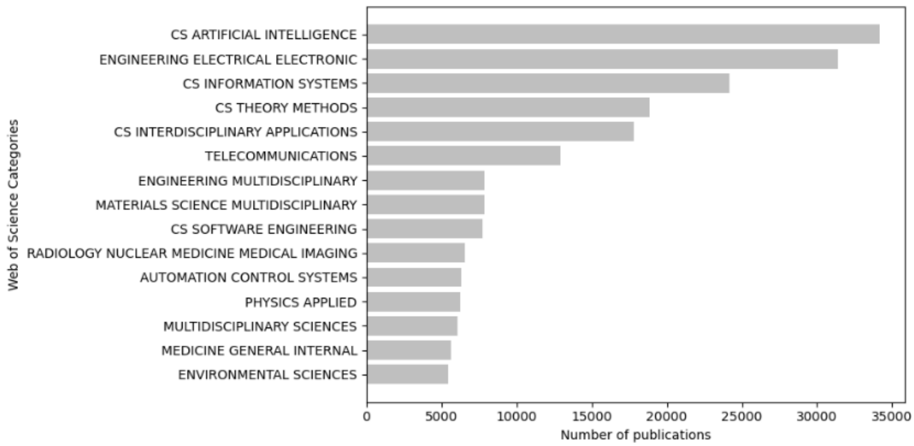

Figure 4. Domains' mapping – WoS AI publication volumes, 1975-2024
Source: Authors' own processing.

As for the conclusions, first of all, we can note four sub-domains of the computer science (CS) field, namely *artificial intelligence*, *information systems, theory methods*, and *interdisciplinary applications.* Besides those, on the second spot, we can also see the *engineering electrical and electronic* (EEE) category, referring to the more technical approaches. Completing the list, we have various complementary fields such as *education*, *telecommunications*, and *medicine,* all of which benefit from the high volume of AI case studies designed to improve their connex practices and processes. At the same time, *environmental sciences* is a category that is entering the top 15, illustrating a strong correlation between the last years' focus on the environmental transition and the broad potential that AI can add to this axis.

Next, after getting a glimpse at the general distribution, we can now move forward to the highlighted hubs. The first view that we can use in order to reach our objective is visible in Figure 5, where we illustrate the differences between the American and Chinese number of papers, registered under different WoS domains. With red bars, we can see the positive differences shown by China (so the categories in which the country published more papers than the USA), while with blue we have the domains in which the USA has exceeded its opponent contributor.

In terms of interpretation, we can assess how China is leading in a vast majority of domains, especially the *EEE* and *computer science* ones. On the other hand, there is a clear distinction in volumes when it comes to the medicine axis for the American state. Fields such as *radiology, cardiovascular systems*, *surgery, health sciences, medical informatics,* or *clinical neurology,* are showing a higher interest coming from the US academia, a fact directly correlated with the advanced healthcare system and substantial financial resources allocated to medical research and innovation across the states.

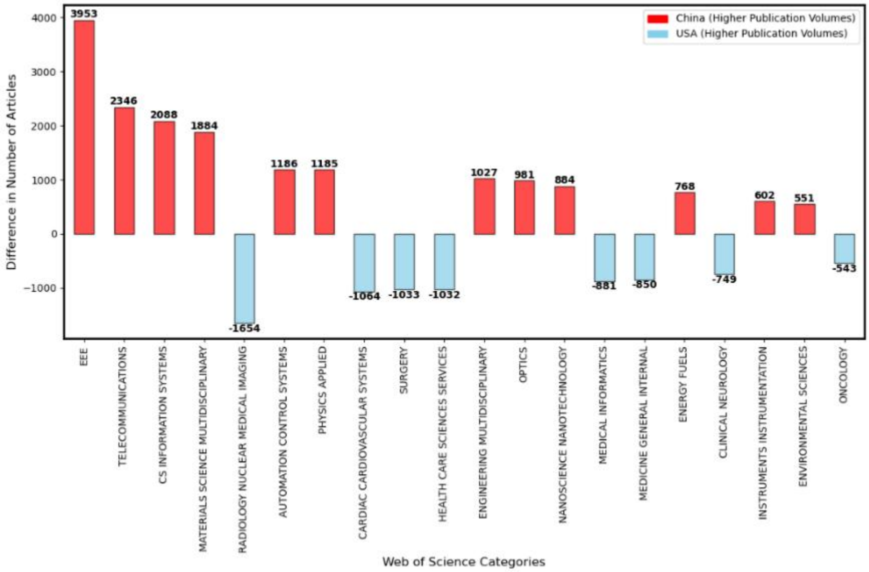

Figure 5. Domains' differences – China and the USA – WoS AI publication volumes, 1975-2024

Source: Authors' own processing.

Thus, as a first conclusion to this objective, China seems to be more focused on the engineering, telecommunication, and automated part of AI applications, while the USA has leveraged its potential throughout the health industry.

VOSViewer. Network and clustering

Proceeding now to a more in-depth analysis, we will finally use VOSViewer application in order to generate a co-occurrences network for both China and the USA. An important methodological mention is that we are going to perform the analysis for the 2024 year, considering it the most relevant for the current status quo. Therefore, for each of the two geographical poles, we will present the network graphical visualisation, followed by some insights into the clustering results and word linkages. As input, we have used the title and abstract for approximately 12,260 publications coming from China and 11,300 publications coming from the USA. As a methodological mention, it is also crucial to underline the approximation part, considering that the numbers on the Web of Science platform are relatively volatile, with some papers still being in different publication steps (so the actual volumes can increase, but not dramatically).

Regarding the further parametrisation of the inputs, we have plotted only the words that have a full count frequency of a minimum of 50 in the entire corpus of documents (1,217), while the number of relevant terms threshold will be set at 60% (730). The relevance score is calculated by VOSViewer, using the *Term Frequency-Inverse Document Frequency Methodology (TF-IDF)*. In this way, we can make sure that the graphs are not overly populated and that they can reflect the correct insights.

Starting with the USA, in Figure 6 we can observe the main co-occurrence graph, as well as the distribution into three main clusters.

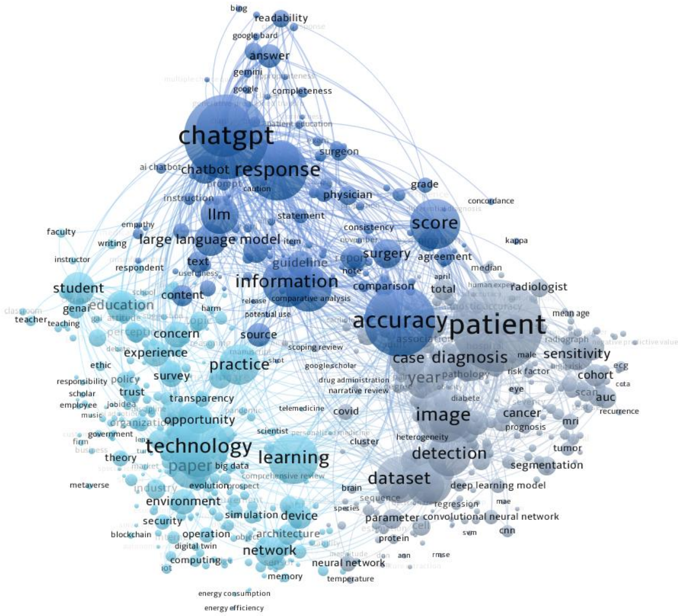

Figure 6. Co-occurrence network – WoS AI publication volumes – USA, 2024
Source: Authors' own processing.

However, in order to better define these sub-groups, we have also illustrated the top most linked words within each cluster (Table 3), making it easier to describe the research axes and to highlight the specific characteristics.

Table 3. Clusters description – WoS AI publication volumes – USA, 2024

Cluster	Cluster Naming	Strongly linked words
Cluster 1	Technology Linkages & Impact	technology, learning, paper, practice, network, student, education, opportunity, device, environment, perspective, experience, concern, future, implication
Cluster 2	Healthcare & Medical Imaging	patient, image, dataset, diagnosis, detection, prediction, disease, feature, case, sensitivity, test, age, classification, specificity, deep learning, cancer
Cluster 3	AI & Language Modelling	ChatGPT, accuracy, response, question, score, information, level, LLM, GPT, large language model, surgery, chatbot, guideline, recommendation

Source: Authors' own processing.

Therefore, the blue (superior) cluster will have as the most used word *ChatGPT*, since it represents the American innovation of the decade. In addition to this, we can notice highly linked AI concepts, such as *accuracy, response, information, score, large language model (LLM)*, or *chatbot*. Therefore, this cluster can be broadly interpreted as *AI & Language modelling.*

Going further, we have the medical cluster (the grey one), an agglomeration that is not surprising considering the previous insight and the emphasis of the US academia on this axis. Words such as *patient, image, diagnosis, detection, disease*, or *sensibility* will be representative of this national priority, while the broad cluster can be defined as *Healthcare & Medical Imaging.*

Last but not least, we have a third cluster, in which we can pinpoint the existence of both technological terms such as *technology, device, simulation, computing, network,* and the intersection with other development or socioeconomic aspects, such as *education, environment, learning,* or *practice*. Generally, this allocation of words can be named *Technology Linkages & Impact.*

After getting a glimpse of the USA distribution, we can now study the second research hub identified, namely China, using the same methodical steps. As seen in the co-occurrence map plotted in Figure 7, the algorithm generated 4 clusters of dominant keywords, with the same criteria being used: words that have a frequency of more than 50 times (1,171) and a relevance score of over 60% (703).

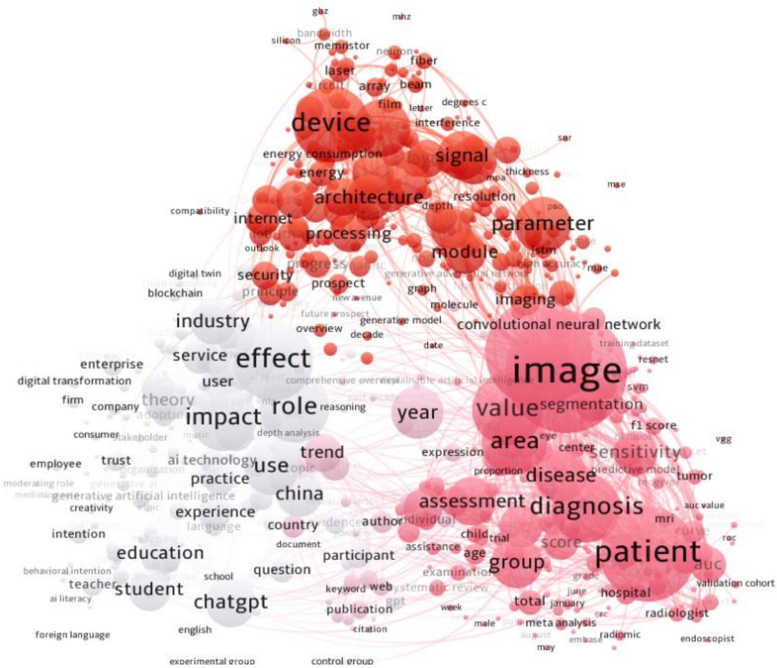

Figure 7. Co-occurrence network – WoS AI publication volumes – China, 2024
Source: Authors' own processing.

Next, the same approach will be undertaken, by looking also at the strongly linked words from each cluster in order to define and name them (Table 4). As a first and general insight, we can note that the main cluster of states will not differentiate as much in terms of general interpretation, having similar word prevalences.

Table 4. Clusters description – WoS AI publication volumes – China, 2024

Cluster	Cluster Naming	Strongly linked words
Cluster 1	Technological advancement in devices and infrastructure	device, parameter, sensor, signal, neural network, architecture, imaging, communication, layer, processing, module, operation, power, internet, mode
Cluster 2	Medical imaging, diagnosis, and engineering	patient, image, feature, value, diagnosis, area, group, disease, sensitivity, AUC, assessment, treatment, classification, curve, outcome, specificity, test, cancer
Cluster 3	The impact of AI tools on connex socioeconomic fields	effect, role, impact, use, ChatGPT, China, student, education, industry, perspective, context, theory, experience, behavior, robot, practice, service
Cluster 4	Research Universe	year, trend, science, country, cell, publication, author, university, web, institution, medicine, topic, covid, bibliometric analysis, expression, journal, pathway

Source: Authors' own processing.

For example, starting with the red cluster, we can notice once again several highly technological terms, such as *device, architecture, signal, parameter, neural network*, and *layer,* concepts that are crucial for the general devices and infrastructure progress. However, as opposed to the USA, this cluster will be a well-defined one, with a clear focus on the Asian country towards technological advancements and highly complex artificial intelligence solutions.

Next, for the second cluster (the grey one), we have a more generic agglomeration of words and concepts. One of them will be the quantitative and applicative one, where notions such as *impact, effect, role, use,* and *experience* will be the most connected, while the education and theoretical dimension is also emphasised through words such as *student, theory, behaviour,* or *practice.* At the same time, an interesting observation comes from the fact that we will not see *ChatGPT* as distinctive as it was in the USA visualisation, a sign that this technology has not penetrated Asian academia as intensely. Considering all these, we will define the current cluster as *The impact of AI tools on the connex socioeconomic axis.*

Going forward, in the hot pink cluster we can profile a clear majority of medical terms, represented by words such as *patient, diagnosis, image, hospital,* or *disease,* along with several AI-specific concepts such as *validation, precision, AUC, curve,* and *test.* This will reflect the strong relationship between medical research activity and AI applications, similar to the case of the USA. Last but not least, regarding the fourth cluster (the light pink in the background), we can distinctively name this one as *The research and academia universe*, with words such as *trend, author, publication, year, university, institution,* and *bibliometric analysis* being the ones with the strongest connections.

Thus, through this second overview, we have managed to confirm the previously mentioned insights, as well as profile the most important dimensions of AI, considering both the similarities and differences between the two.

In order to complete the exploration, we propose one last view, consisting of the most relevant and frequent terms in each country, in a comparative approach. Therefore, in Figure 8 we have plotted the words that occur more in the USA publications, while in Figure 9 we have the reversed perspective, featuring the more frequent concepts for China.

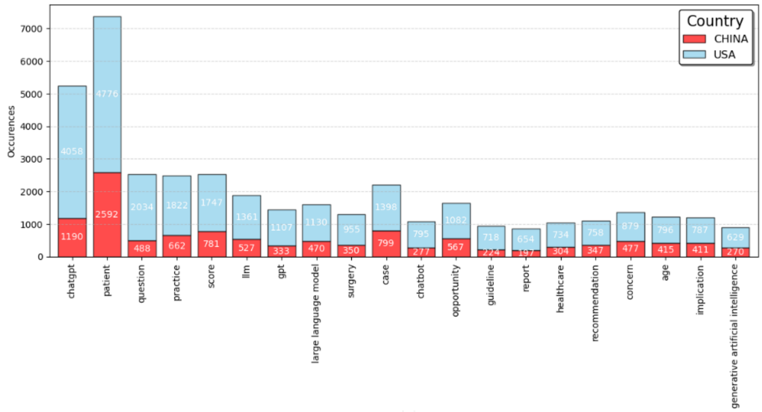

Figure 8. Highest differences in concepts – WoS AI publication volumes – USA, 2024
Source: Authors' own processing.

Thus, in the case of the USA, elements such as *ChatGPT*, *GPT*, *generative artificial intelligence, LLM,* or *large language models* have been much more used than in the case of China, profiling a clearly developed AI priority and a correlated research interest throughout 2024. At the same time, confirming the health precedence, words such as *patient, surgery, case,* or *healthcare* are also more common in the case of the American research community.

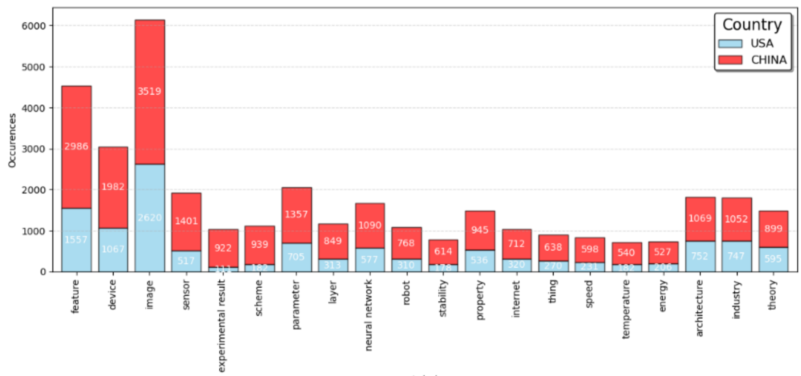

Figure 9. Highest differences in concepts – WoS AI publication volumes – China, 2024
Source: Authors' own processing.

On the other hand, for China, we can see once again the prevalence of highly technological words, such as *device, image, sensor,* and *robot.* At the same time, *neural networks, layers,* or *parameters,* mark an algorithmic preference for deep-learning applications, most likely in dimensions such as *industry* or *energy*.

Finally, we consider the third objective reached, with several clear specificities described for each of the two research hubs.

5. Conclusions

Over the exploration of our subject, we have seen how a bibliometric analysis could prove to be a valuable instrument for understanding the past, current, and future perspectives of artificial intelligence. Since research usually follows the technological and computing progress coming from private actors, the growth in the publication volumes with AI as a topic was inevitable, being the most evident starting from 2010 and then exploding throughout the 2020s. As for the future predictions, using an ARIMA model, we were able to generate a forecasted number of academic papers by 2027, doubled versus the 2024 volume, a fact, however, that needs to be interpreted with prudence.

Going next, we have explored the geographical distribution of AI research, with two major poles being prevalent: the USA and China, and several other macro contributors following them. In general, the domain seems to have gained popularity across different geographical poles, with Europe, North America, and Asia being prevalent.

Followingly, we proceeded to profile the specifics of the two leading hubs, this time considering the most recent year of publications: 2024. Thus, we have generated insights into both similarities, coming mainly from the general comparable topic clustering, as well as country specificities. Between these we can mention a more pronounced interest in large language models, generative AI, and AI-enhanced medical applications coming from the USA, while China will lead the race when it comes to technology, automation, robots, or deep learning.

Overall, we have seen how the evolving AI research ecosystem reflects both rapid innovation and regional specificities. As AI continues to shape industries and societies, future studies should monitor these trends to better understand the field's trajectory and its global impact.

Acknowledgements: *This paper was co-financed by The Bucharest University of Economic Studies during the PhD program.*

References

[1] Dissanayake, H., Manta, O., Iddagoda, A., Palazzo, M. (2024), *AI applications in business: Trends and insights using bibliometric analysis. The International Journal of Management Education*, 22(3), 101075.

[2] Gao, F., Jia, X., Zhao, Z., Chen, C.C., Xu, F., Geng, Z., Song, X. (2021), *Bibliometric analysis on tendency and topics of artificial intelligence over last decade. Microsystem Technologies*, 27, 1545-1557.

[3] Gao, H., Ding, X. (2022), *The research landscape on the artificial intelligence: a bibliometric analysis of recent 20 years. Multimedia Tools and Applications*, 81(9), 12973-13001.

[4] Guan, C., Mou, J., Jiang, Z. (2020), *Artificial intelligence innovation in education: A twenty-year data-driven historical analysis.International Journal of Innovation Studies*, 4(4), 134-147.

[5] Hajkowicz, S., Sanderson, C., Karimi, S., Bratanova, A., Naughtin, C. (2023), *Artificial intelligence adoption in the physical sciences, natural sciences, life sciences, social sciences and the arts and humanities: A bibliometric analysis of research publications from 1960-2021. Technology in Society*, 74, 102260.

[6] Hinton, G.E., Rumelhart, D.E., Williams, R.J. (1986), *Learning representations by backpropagating errors. Nature,* 323(6088), 533-536, https://doi.org/10.1038/323533a0.

[7] Ho, Y.S., Wang, M.H. (2020), *A bibliometric analysis of artificial intelligence publications from 1991 to 2018. COLLNET Journal of Scientometrics and Information Management*, 14(2), 369-392.

[8] Javidan, A.P., Li, A., Lee, M.H., Forbes, T.L., Naji, F. (2022), *A systematic review and bibliometric analysis of applications of artificial intelligence and machine learning in vascular surgery. Annals of Vascular Surgery*, 85, 395-405.

[9] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998), *Gradient-based learning applied to document recognition. Proceedings of the IEEE,* 86(11), 2278-2324. https://doi.org/10.1109/5.726791.

[10] Liu, L., Duffy, V.G. (2023), *Exploring the future development of Artificial Intelligence (AI) applications in chatbots: a bibliometric analysis. International Journal of Social Robotics*, 15(5), 703-716.

[11] Maisonobe, M. (2022), *The future of urban models in the Big Data and AI era: a bibliometric analysis* (2000–2019). *AI & Society*, 1-18.

[12] Marinas, L.E., Paun, C.V., Diaconescu, M., Smirna, T.G. (2024), *Artificial Intelligence Readiness and Employment: A Global Panel Analysis. Economic Computation and Economic Cybernetics Studies and Research*, 58(4), 57-74.

[13] McCarthy, J., Minsky, M., Rochester, N., Shannon, C.E. (1956), *A proposal for the Dartmouth summer research project on artificial intelligence.* Dartmouth College.

[14] Minsky, M. (1988), *The society of mind*. Simon & Schuster LLC, California, USA.

[15] Newell, A., Simon, H.A. (1972). *Human problem solving*. Prentice-Hall.

[16] Pearl, J. (1988), *Probabilistic reasoning in intelligent systems*. San Mateo, CA: Kaufmann, 23, 33-34.

[17] Riahi, Y., Saikouk, T., Gunasekaran, A., Badraoui, I. (2021), *Artificial intelligence applications in supply chain: A descriptive bibliometric analysis and future research directions. Expert Systems with Applications*, 173, 114702.

[18] Shukla, A.K., Janmaijaya, M., Abraham, A., Muhuri, P.K. (2019), *Engineering applications of artificial intelligence: A bibliometric analysis of 30 years (1988–2018). Engineering Applications of Artificial Intelligence*, 85, 517-532.

[19] Tran, B.X., Vu, G.T., Ha, G.H., Vuong, Q.H., Ho, M.T., Vuong, T.T., ..., Ho, R.C. (2019), *Global evolution of research in artificial intelligence in health and medicine: a bibliometric study. Journal of Clinical Medicine*, 8(3), 360.

[20] Turing, A.M. (1950), *Computing machinery and intelligence. Mind,* 59(236), 433-460, https://doi.org/10.1093/mind/LIX.236.433.

[21] Xie, B., Xu, D., Zou, X.Q., Lu, M.J., Peng, X.L., Wen, X.J. (2024), *Artificial intelligence in dentistry: a bibliometric analysis from 2000 to 2023. Journal of Dental Sciences*, 19(3), 1722-1733.