Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

Volume 06 Issue 02
2024

Research Article
https://doi.org/10.56979/602/2024

Analyzing Paper Citation Trend of Popular Research Fields

Lubna Zafar[1], Nayyer Masood[2], Fazle Hadi[3], and Sheeraz Ahmed[4*]

[1]University of Poonch Rawalakot, 12350, Azad Kashmir, Pakistan.
[2]Capital University of Science and Technology, Islamabad, 44000, Pakistan.
[3]Higher Education Department, Khyber Pakhtunkhwa, 19110, Pakistan.
[4]Iqra National University, Peshawar, 25000, Pakistan.
[*]Corresponding Author: Sheeraz Ahmed. Email: sheeraz.ahmad@inu.edu.pk

Received: January 09, 2024 **Accepted:** February 25, 2024 **Published:** March 01, 2024

Abstract: The ever-expanding volume and diversity of scientific literature pose a significant challenge for researchers in detecting emerging, current, and future research trends. A trend represents the prevailing direction of research within a defined timeframe. Detecting trends involves identifying areas of growing interest over time, while trend analysis involves gathering data and discerning patterns. Despite the utilization of diverse methods for analyzing and identifying trends in scientific research, there remains a lack of comprehensive understanding regarding the significance of following research trends for citation of research papers. The objective of this research is to examine the significance of monitoring trends in Computer Science (CS) research, the influence of aligning with these trends on paper citations, and the correlation in citation patterns among papers within the CS domain. We analyze trends in CS conference papers and the evolution of research fields from 1985 to 2017 using the Microsoft Academic Graph (MAG) dataset of CS papers in the L1 field of study (FoS). Our experimental findings reveal that Data Mining, Artificial Intelligence, Computer Vision, Machine Learning, and Database research exhibit the highest publication trends. Additionally, our results suggest that papers within the same field demonstrate similar citation trends.

Keywords: Computer Science; Trend; Citation Trend; Field of Study Trend.

1. Introduction

The exponential expansion of social networking sites like Facebook and Twitter has resulted in a significant increase in the accessibility of information disseminated through these channels [1]. People are increasingly depending on these platforms as their main sources of information, although there's frequently minimal supervision regarding the credibility of the content being produced, leading to the rampant proliferation of rumors [2], [3].

The quantity and variety of scientific literature are continuously expanding, with nearly 2.5 million new research articles published each year, as indicated by a study [1]. This presents a considerable challenge for researchers in keeping up with emerging, current, and forthcoming research trends. A trend refers to the overall direction in which research is progressing within a specified timeframe and is characterized as an area that gains significance and effectiveness over time. Detecting trends involves identifying subject areas that are experiencing increasing interest over time. Trend analysis entails gathering information and attempting to identify patterns within data [2]. A significant challenge in trend detection and analysis lies in recognizing research trends within scientific research. Facilitating the detection and analysis of research trends streamlines this process, enabling researchers to promptly identify emerging trends in scientific topics and explore the most recent related subjects within their research domain.

In recent years, notable advancements have been made in techniques for detecting and analyzing trends. Initially, research trends were primarily identified by domain experts who were selected for their extensive knowledge and experience. Even today, the process of consulting experts remains widely utilized in the field of science and technology policymaking. Nevertheless, manually identifying research trends requires substantial labor and time due to the continual expansion of research literature an-nually. Hence, there is a crucial need to strategize, enhance, and automate this process [3].

Detecting and understanding emerging trends and swift shifts within scientific fields can significantly enhance researchers' ability to adapt to changes promptly. Trend detection reflects the scientific research areas garnering the most attention from researchers. The combined focus frequently highlights the main issues to address or emerging topics of promise within each field, offering policymakers valuable insights. Tracking research trends aids in resource allocation and technology prediction [4], underscoring the significant implications of trend detection in research.

The field of research is constantly evolving, constantly evolving and expanding, with new and innovative topics emerging while others fade into obscurity. These fluctuations present a precise challenge. It is crucial to identify significant innovative research trends and forecast their future impact, not only for established stakeholders such as researchers, academic publishers, official funding organizations, and companies operating in cutting-edge industries, but also for those whose survival and success depend on staying at the forefront of innovation. To achieve this, there is a growing demand for specialists and tools capable of detecting, comprehending, and predicting research trends.

Figure 1 illustrates the trend of keywords in computer science papers from 1990 to 2015 on the x-axis, with the score of keywords indicated on the y-axis [5]. This score is used to gauge the level of importance of specific research topics or areas. The figure highlights keywords with the highest scores, indicating their popularity and growth over time. Until 2012, web search engines were dominant, after which social relations and recommender systems gained prominence.

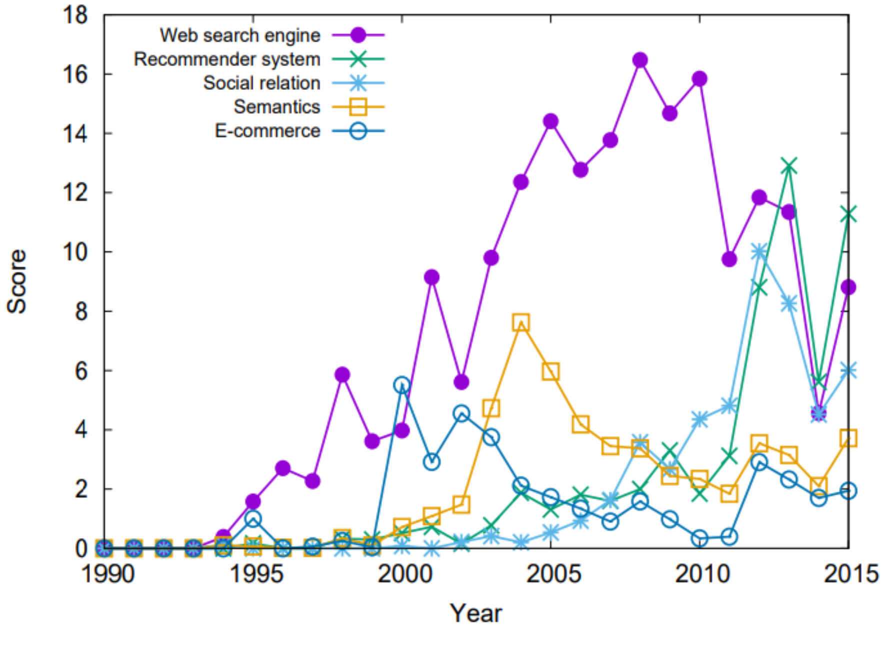

Figure 1. Computer Science keyword-based trend

Trend detection often involves utilizing words or phrases extracted from bibliographic data found in research papers, including the paper's title, abstract, and keywords. This data is regarded as offering a reasonably thorough portrayal of the paper's subject [6] and is consistently employed to quantitatively evaluate research trends [7, 8, 9, 10]. However, the specific selection methods for detecting trends vary depending on the approach to data processing.

Author-selected keywords and Keywords Plus are commonly selected for analysis, although there is limited research demonstrating the efficacy of Keywords Plus [5, 11]. Author Keywords consist of terms believed by authors to best represent their research content, while Keywords Plus include terms frequently found in the titles of a paper's references but not necessarily in the article's title or as Author Keywords. Keywords Plus are generated through automated computer processes [6, 12].

According to [6], Keywords Plus terms are capable of effectively capturing the essence and diversity of a paper's content. The concept of emerging co-word maps for studying semantic associations in scientific literature was introduced and supported by [13, 14], [15, 16]. Co-word analysis is a methodology actively utilized for assessing the strength of relationships among information items in textual data, as shown in Figure 3. This approach directly identifies sets of keywords shared by publications, thereby delineating the themes of scientific literature based on keyword collaborations [8, 17]. Co-word analysis operates on the premise that keywords provide a suitable representation of a paper's content, as depicted in Figure 2. Hence, keywords can be employed to represent the structure of a research field's content.Top of Form

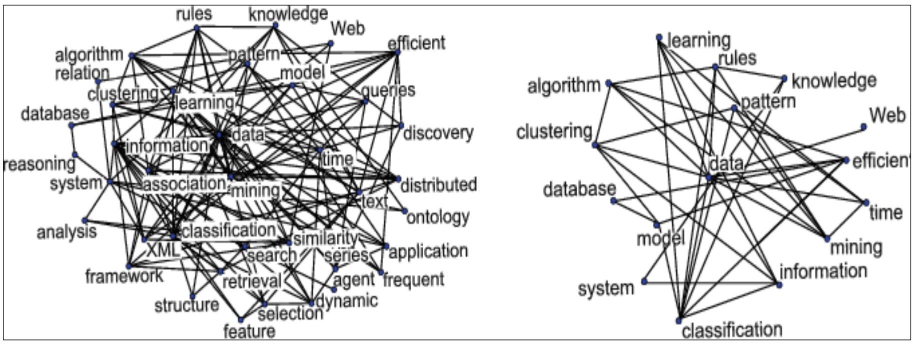

Figure 2. Co-occurrence analysis of keywords

Originating as a recognized research field in the 1960s [18], Computer Science (CS) has undergone rapid advancements and evolved into a mature discipline. A wealth of insights into the developments and trends within Computer Science can be gleaned from analyzing publication metadata. Research within Computer Science progresses dynamically, with new areas emerging and older ones experiencing varying levels of prominence, and sometimes even witnessing a resurgence [19]. The analysis of CS trends using IEEE and ACM research papers is explored in [20], while [21] utilizes the Microsoft Academic Graph (MAG) dataset for trend detection in CS research papers.

This study delves into trends observed in CS conference papers and tracks the evolution of research areas within CS from the years 1985 to 2017. It sheds light on the importance of monitoring trends in research within the field of CS , the impact of trends on citations received by research papers, and the influence of trends on author ranking. Additionally, it aims to differentiate between trend innovators and trend followers. As far as we know, there has been no investigation into the impacts of trends on papers in Computer Science, nor has there been an analysis of the citation accumulation over time for papers that adhere to trends. This study intends to investigate the repercussions of aligning with popular trends on the careers of scientific authors.

In scientific literature, staying attuned to research trends and dynamics can yield notable advantages, aligning with the current interests and aspirations of researchers. Being "research trendy" doesn't imply adherence to traditionalism or conforming to a predetermined plan; rather, it involves a dynamic awareness of innovative insights. Researchers who remain updated tend to appear more contemporary compared to those who neglect to follow evolving trends. Keeping abreast of the latest research trends offers researchers a competitive edge in a rapidly evolving landscape, facilitating better professional connections and enhancing the impact of their papers through increased citations and recognition such as MIPs. Therefore, it is essential to stay informed about research trends in scientific inquiry, as trending issues can shed light on researchers' preferences, objectives, and the latest developments.

Furthermore, it is widely acknowledged that editors play a crucial role in identifying emerging topics and publishing content that is both relevant and motivating. Editors often have the foresight to recognize the significance of emerging trends and are instrumental in bringing attention to these issues through publication. Similarly, funding agencies recognize the impact of research trends on scientific literature and may strategically analyze the evolution of a field to inform their funding decisions and prioritize research areas. This understanding underscores the importance of ongoing inquiry for researchers, editors, and funding agencies, enabling them to track the progress of research fields with the passage of time. The main aim of this study is to investigate the importance of research trends in Computer Science by assessing their influence on scientific research articles.

Our objective is to examine the development of research trends in the field of Computer Science from 1985 to 2017 and categorize periods of trendiness. This study holds potential benefits for researchers, subject experts, and policymakers alike. Researchers and subject experts can quickly discern which trends have dominated their field and identify areas that have received less attention during the specified timeframe. The findings of this study could assist policymakers in allocating research funding to specific topics and fields with greater confidence.

The structure of the paper is outlined as: Section 2 presents related work. Sections 3-4 delve into the significance of the research, the problem statement, and the research questions. Section 5 explains the proposed methodology, while observations and potential future directions discusses in Section 6.

2. Literature Review and Background Study

Various studies in the literature have delved into the problem of trend detection, which has garnered significant interest and practical application across a spectrum of systems including blogs [22, 23], emails [24], social networks [25, 26], and scientific data [4, 27, 28, 29, 30, 31, 32]. Trend detection and analysis methodologies found in literature can be broadly categorized into Probabilistic Topic Models, Citation Networks, and Keyword-based techniques.

Hierarchical LDA [23] has been developed, wherein topics are organized in a hierarchical manner. Advancements have integrated different types of research data, such as semantic topic models [27], facilitating trend identification in new research domains by describing topics and the methods utilized for author identification. Research papers, encompassing exploration, experimentation, evaluation, or argument within specific disciplines, are typically disseminated in venues like conferences, journals, or workshops [30], with LDA being employed to detect emerging topics. Selecting the appropriate publication venue presents challenges for authors, necessitating consideration of various factors including target audience, subject area, and venue reputation.

Illustrated by [32], a semantic topic model presented an algorithm named Klink, adept at discerning relationships and building a semantic network of research domains from a corpus of papers, utilizing heuristic rules, statistical techniques, and external knowledge. An author presented the correlated topic approach[33], which employs the logistic dispersal as an alternative to LDA, aims to overcome LDA's constraint in capturing connections among diverse components. Latent Dirichlet Allocation (LDA) [34], widely utilized for extracting distinct topics from data, has been extended and adapted for various applications.

A Finite Mixture Model proposed in [35] illustrates the association of research areas and examines variations over time to detect emergent research topics. However, its evaluation primarily relies on electronic data, raising questions about its efficacy on research data. Authors propose a semantic topic approach to elucidate the birth of new topics, acknowledging the emergence of alternative types such as online blogs or repositories, increasingly utilized within research communities [36].

A prevalent challenge in these methodologies lies in the complexity of discerning research domains from the generated topic models. A comparable method was introduced in probabilistic Latent Semantic Analysis (pLSA) by [37]. In this approach, words from papers are represented as samples from a fusion model. Utilizing the Expectation-Maximization algorithm, it identifies topics characterized as multinomial random variables, which constitute the fusion's components. Another illustration involves a semantic topic model where topics are linked via semantic connections, constructing a semantic network of research domains.

A semantic classification of research subjects is introduced [38], yielding better results than keywords for discovering research groups. Some methods address this issue by relying on topic classifications. Although these methods build on keywords, they add a theoretical layer to mitigate some drawbacks, eliminating keywords that don't represent a topic well and establishing connections among keywords.

The scientific community has proposed various methodologies for detecting trends beyond LDA. For example, citation networks have been suggested for organizing documents temporally to uncover topic evolution and emerging trends, utilizing citations to assign weights to key terms in papers. Similarly, citation networks employ ACM Digital Library classification to examine the advancement of graph areas and observe trends in scientific research. However, classifications created by humans tend to evolve gradually, and in rapidly changing fields like Computer Science, it's essential to rely on continuously updated classifications. Research papers are categorized [41] based on significant words from titles and abstracts, investigating variations in publications associated with these topics using a citation network.

Nevertheless, as observed [42] in citation network papers, keywords are not preprocessed and may not adequately reflect the significance of research topic areas in different contexts. Additionally, different keywords may address similar topics, making it challenging to discern trends. A network of co-occurring keywords [43] in scientific data is utilized to identify trends and emergent research topic areas over time. Techniques such as patent analysis, bibliometric study, and text-mining analysis [44] are employed to identify research trends.

A proposed method [45] compares the dispersion of keywords extracted from research data using citation graphs associated with publications containing these keywords. This technique assumes that if a keyword term is suitable for a topic area, then the research papers containing the keyword will have stronger links than normal ones. However, this technique may not be suitable for areas in early stages of development.

A widely used technique involves employing keywords to represent research topics, known as a keyword-based topic model. This method is commonly utilized for analyzing research trends [3], although it can introduce bias towards papers with multiple Field of Study (FoS) classifications. The correlated neural influence model [46] delves into research evolution within conferences and reveals connections among different conferences. While this concept has yet to be applied to unveil influence mechanisms among scientific conferences, a two-dimensional text mining approach [40], which includes clustering and bibliometric analysis, has been proposed. Bibliometric analysis of keywords [48] is employed to examine the knowledge structure of scientific research in journals. Techniques like Saffron and MAS [5] also make use of keyword topic model methods. However, this approach encounters several challenges. Primarily, keywords may lack consistency and could include terms that do not represent topics, such as "case study." Furthermore, topics may exhibit a hierarchical structure based on macro areas with specific sub-areas, which is not accounted for by the keyword topic model, where relations between research topics are not explicitly defined [10].

Furthermore, the interpretation of similar keywords can vary significantly. A keyword-based topic model often struggles to capture synonyms, resulting in keywords representing the same concept (e.g., "ontology-based," "ontology," "ontologies") being treated as distinct topics. These challenges can be addressed by encouraging authors to utilize keywords from the current Association for Computing Machinery (ACM) classification. One approach, as outlined in [18], constructs a network of keywords and then conducts statistical analysis by calculating metrics such as degree, strength, clustering coefficient, and end-point degree to form clusters corresponding to research topics. Another method, as described in [20], utilizes the ACM classification as a means of categorizing subjects and visualizes trends over time based on ACM data.

Another technique, as elaborated in [25], establishes connections between papers and topics using keywords and words extracted from abstracts to analyze topic trends across various time scales. To effectively identify topic trends, it is beneficial to define a topic state based on features such as the number of associated publications/citations [39, 46] and the number of active authors in the field [50], and then observe their evolution over time. Additionally, relational topic modeling, which integrates network structure and Latent Dirichlet Allocation (LDA) of papers to model topic areas, employs citation networks and LDA [21] to address the issue of topic evolution. This method identifies topics in autonomous subsets of data and utilizes citations to link topics across different time periods.

Moreover, a hybrid method [38] is utilized for detecting the evolution of research topics, although it may not adequately capture nascent trends in research. This strategy integrates Probabilistic Latent Semantic Analysis (PLSA) to model topics in a sliding window configuration, enabling the examination of topic progression across time.

Existing research presents a variety of approaches for uncovering research trends but may be limited to previously identified areas. Various techniques are utilized to detect emerging, current, and future research trends from scientific data. However, there remains a lack of comprehensive solutions for analyzing and evaluating the impact of trend following and the significance of adhering to a research trend in scientific research. Questions such as how papers following trends accrue citations over time, how to distinguish between trend innovators and followers, the impact of trends on author rankings, and the effect of adhering to popular trends on the careers of scientific authors remain open and intriguing areas for exploration.

3. Research Methodology

3.1 Dataset Collection

We utilized a dataset sourced from the Microsoft Academic Graph (MAG), which was initially released by Microsoft Research in 2015 [20]. MAG contains a wide array of information concerning publications, including paper ID, title, authors, abstract, keywords, field of study, publisher, and year of publication. Covering papers from various disciplines, MAG provides comprehensive statistics as outlined in Tables 1 and 2. We opted for MAG due to its well-organized structure and the availability of integrated research tools, facilitating efficient data access, analysis, and processing. The proposed methodology is elucidated in Figure 3.

Table 1. MAG multidisciplinary entities

Multidisciplinary Entities	
Entity	Entity Count
Papers	210,776,664
Authors	254,664,479
Journals	48,665
Conferences	4,343
Field of Study	229,793

Table 2. MAG Computer Science entities

Multidisciplinary Entities	
Entity	Entity Count
Papers	1,354,603
Authors	2,324,591
Conferences	1,277
Field of Study	9,800

3.2 Computer Science Papers Extraction Process

The primary focus of this study lies in computer science (CS) papers, and one strategy to identify them is by specifically targeting papers published in CS-centric venues. Historically, distinguishing such venues for CS research has proven challenging, as significant findings are often initially disseminated through conference proceedings rather than journals.

Consequently, in this research, we construct our CS paper dataset from the MAG dataset by filtering for papers that are both categorized under the field of study of CS and exclusively published in conference proceedings. Within the MAG dataset, each paper is associated with various keywords, serving as links to its respective field of study. However, these keywords lack unique identifiers and simply serve as connections between the paper and its field of study.

As a result, we consider each paper as contributing to a set of fields of study within the MAG dataset, with each field representing a distinct research area or topic. Microsoft Academic Graph offers an exploration feature that categorizes the research topics of papers into specific fields of study. By leveraging this feature, we can roughly discern the topic of each paper without the need to delve into its abstract or content [19,20]. Furthermore, we rely on the research trends of CS papers spanning the period from 1985 to 2017, which are stored in the database. This underscores the significance of these research topics within the field of CS.

3.3 Field of Study (FoS)

The process of discovering entities within the field of study involves two main approaches for organizing the data: (1) entities currently designated as field of study categories in the knowledge base, and (2) entities identified through name-matching with keyword features in article entities, as depicted in Figure 5. Previously, the development of Bing's in-house knowledge base, a web search engine operated by Microsoft, played a pivotal role in associating entities. This knowledge base is constructed based on hyperlinks, web-click signals, and entity contents to classify new nominees for fields of study.

The Microsoft Academic Graph (MAG) categorizes fields of study (FoS) in a hierarchical structure spanning four levels, from level0 to level3. Level3 offers the highest level of detail, encompassing 47,989 distinct fields of study, followed by 1,966 at level2, 293 at level1, and 18 at level0. Among a total of

166,192,182 papers, approximately 41,739,531 (about 33%) are associated with one or more FoS entities [28]. Figure 4 depicts the distribution of papers across the 18 level0 FoS.

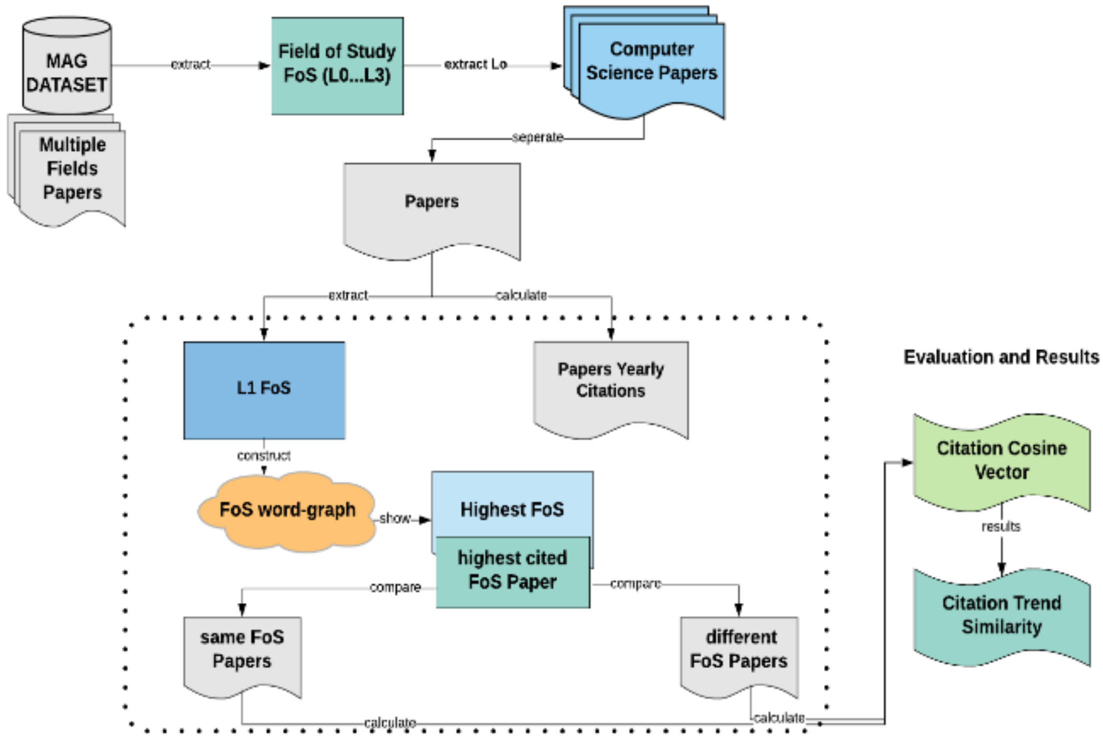

Figure 3. Proposed Methodology

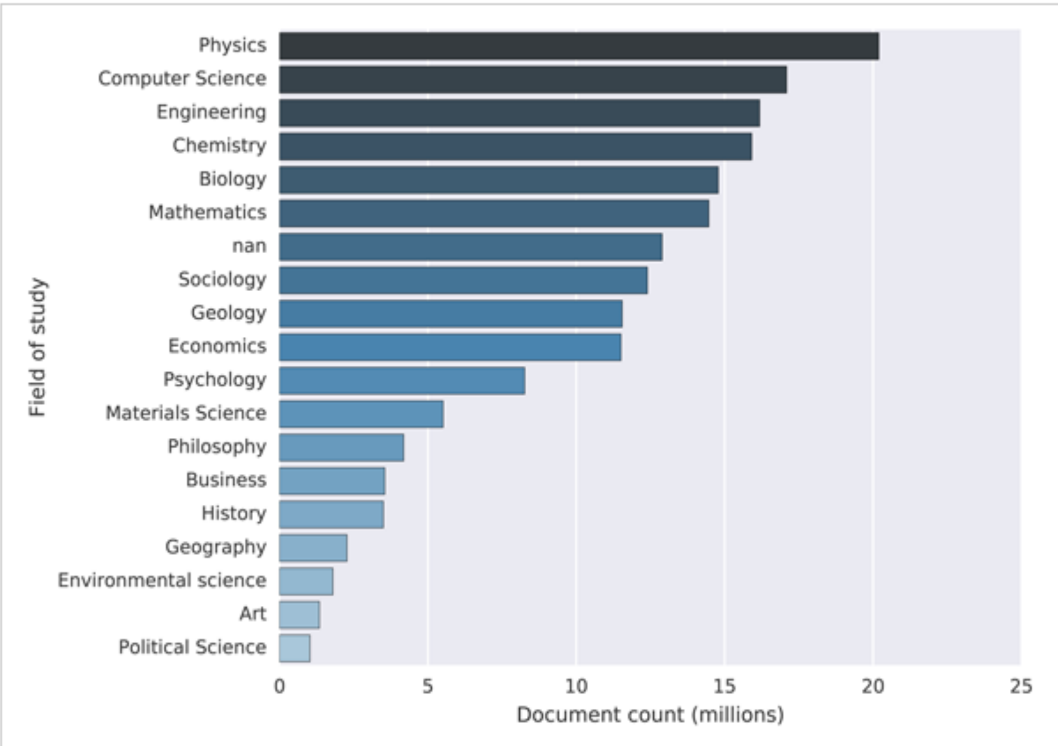

Figure 4. Distribution of papers in MAG

3.4 FoS Graph Construction

We employed a Graph-Based Approach utilizing the Degree Centrality measure for trend detection, employing keywords as depicted in Figure 2. Initially, a keyword graph was constructed from the keywords extracted from Computer Science papers.

Graph centrality measures were employed to assess the significance of each keyword, and papers were subsequently ranked according to the keywords they incorporate. To standardize keywords, we followed a basic principle of normalization. Keywords with synonymous meanings, such as "www" and "world wide web," were unified into one term [29].

In the construction of the keyword graph, a graph G=(V,E) was formed, consisting of a set of vertices or nodes V and E is a set of edges. Each keyword contained within a research paper was represented by a labeled node. The edges were designed to capture the relationships between keywords as they co-occurred within the research papers, as illustrated in Figure 5. The proximity between keywords was represented by the edges connecting the nodes, defined through a specific range of keywords. Keywords forming a graph structure within a paper were considered, while those not forming such a structure were disregarded, as we deemed them less influential based on their occurrences in papers. Once the keyword graph was constructed, the Degree Centrality measure was applied. Scores were computed for each node using equations 1-2-3-4 [25].

Degree centrality quantifies the number of edges linked to a node. In the context of a keyword graph, the degree of a node Vi signifies how many other keywords are associated with the keywords represented by Vi. Denoted as D(Vi), the degree centrality of a node Vi is calculated according to the methodology outlined in reference [25].

$$C_D(V_i) = Deg(V_i) \tag{1}$$

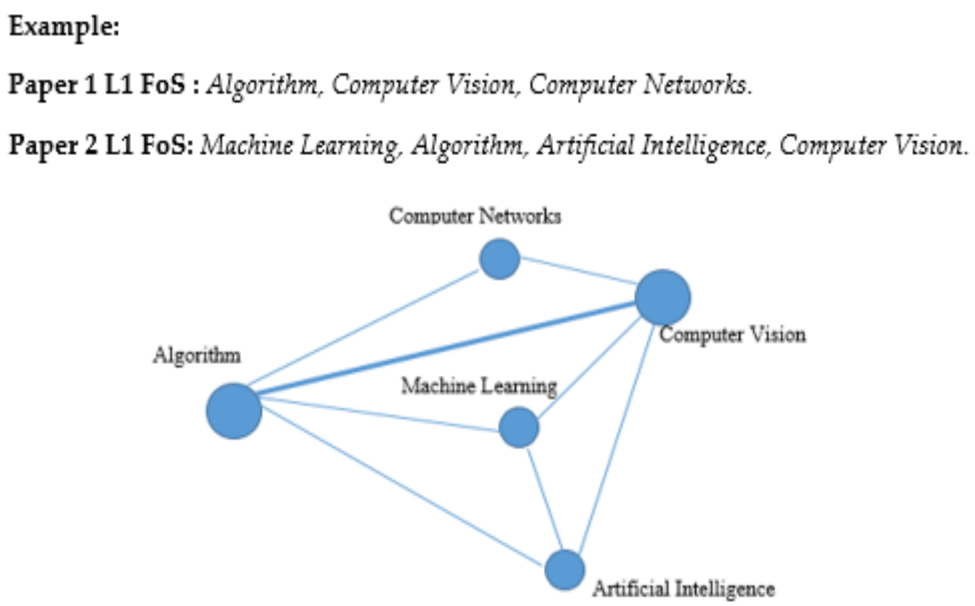

Figure 5. FoS Graph Construction

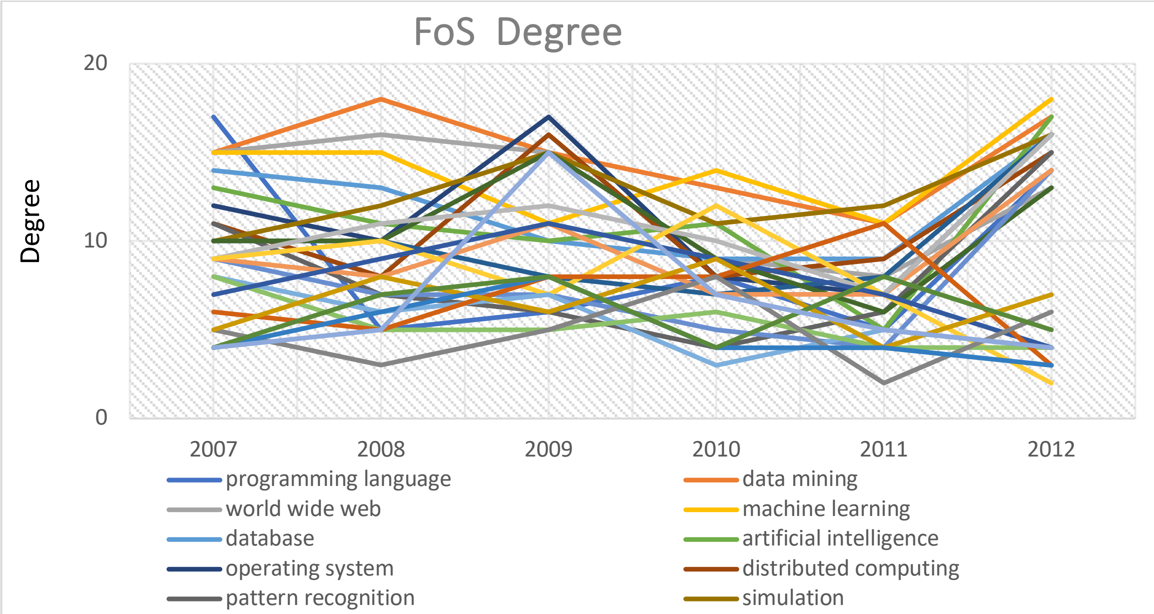

Figure 6. FoS Degree from 2007-2011

3.5 Trend Effect of Paper Citation

Citation analysis, a method in bibliometrics is utilized to evaluate the citations received by publications. This method is frequently employed by researchers to gauge the impact of their work within their respective fields. By quantifying the frequency of citations a paper receives in other research publications, researchers can glean insights into its influence on the discipline. A paper with a high citation count may indicate that it has sparked discussion or critique within its field. Additionally, examining the list of papers citing a particular paper, known as citation chasing, can provide further information on the research topic.

In our study, we utilized research trends in computer science papers from the period 1985-2017, as documented in references [19,20], which are stored in our database. This underscores the significance of the research topics within the field of CS. Subsequently, we extracted papers with the highest citation counts from the same time frame. Firstly, we selected papers containing the first identical keyword and compiled them into a file, as illustrated in Figure 3. Secondly, we identified papers containing common keywords and segregated them into a separate file, as depicted in Table 4. Algorithm 1 outlines the algorithm employed for this process.

Table 3. Papers with same FoS

Papers	Keywords
Paper 1	Sensor network, structural health monitoring, widen, reliability, design
Paper 2	Sensor network, localization, mobility, location-awareness, tracking
Paper 3	Sensor network, energy conservation, tracking

Table 4. Papers with different FoS

Papers	Keywords
Paper 1	Sybil attack, sensor network, security, algorithms.
Paper 2	Secure localization, sensor network, range independent
Paper 3	Medium access control, sensor networks, wireless network, energy efficiency

Algorithm 1. Algorithm for citation trend of papers

```
Step 01: Start
Step 02: Input: Papers P [P1,P2,….,Pn]
Step 03: Output: Cosine trend similarity
Step 04: Separate P of first keyword same and common keyword
              P={(P1=a,b,c),(P2=a,d,e)}      //first keyword same
              P={(P1=a,b,c),(P2=d,c,e)}     // common keyword same
Step 05: Calculate yearly citations of each P
               PCitperyear =PCitedperyear    //Cit: citation
Step 06: Create document matrix for the collection
Step 07: if (Y1 citation < Y2 citation)    // Y1:Year1,Y2:Year2
                          Y1 = 1
              else
                           Y1 = 0
Step 09: Calculate citation cosine vector similarity of highest cited paper with P1,P2,…..,Pn
Step 10: Calculate average
Step 11: Stop
```

4. Results and Discussions

4.1 Evaluation Metric

The procedure utilizes the Citation Cosine Vector Similarity Metric to discern similar trends in citations among papers. This method involves comparing documents by computing the cosine similarity between their citation term vectors [54]. A high cosine similarity value indicates similarity between documents, with values of "i" and "j" being close and the angle between vectors being small. In our analysis, we compute the cosine vector similarity for papers within the same Field of Study (FoS) and those in different FoS categories, as illustrated in Figure 6.

The cosine vector similarity measure can be described as;

$$Cos\ (\emptyset) = \frac{i.j}{||i||.||j||} \tag{2}$$

2007 cites	2008 cites	2009 cites	2010 cites	2011 cites	2012 cites	2013 cites	2014 cites	2015 cites	2016 cites	Cit.Cos.Vec
-50	-50	-50	0	-30	-20	10	60	30	10	1
-23	-5	-50	0	-10	-20	0	-20	50	60	0.97766682
-8	-5	-45	-20	-20	10	-10	20	-10	60	0.95568914
-12	-8	-11	-38	-4	-20	30	4	10	-14	0.97087272
-5	-1	-2	-37	-16	6	-10	-10	-10	60	0.88710537
										0.94355269

Figure 7. Cosine vector similarity of same FoS papers

Figure 7, presents papers within the same Field of Study (FoS), where "Sensor network" stands out as the primary keyword. Additionally, the table includes yearly citation counts for papers spanning from 2004 to 2017, reflecting the total citation count over the specified time period. The citation vector similarity matrix displayed in Table 6 illustrates citation vectors, with a 1 marking indicating that the citation count of one year is less than the subsequent year, and 0 otherwise. Subsequently, we assess the citation cosine vector similarity of the paper with highest citation count with other papers within the same FoS, as well as those from different FoS categories, as depicted in Table 7. Finally, we compute the average of the citation cosine vector similarity values. Preliminary findings suggest that papers sharing the same FoS exhibit a similar citation trend compared to those from different FoS categories.

2007 cites	2008 cites	2009 cites	2010 cites	2011 cites	2012 cites	2013 cites	2014 cites	2015 cites	2016 cites	Cit.Cos.Vec
-50	-50	-50	0	-30	-20	10	60	30	10	1
-38	-130	-100	-60	40	-90	-70	-20	-30	30	0.94666563
-8	16	-94	-60	-70	-50	-20	-260	90	200	0.80971897
-16	-60	-30	-60	80	0	-200	30	-90	-40	0.84339592
-28	-90	-150	-20	-60	0	-20	70	20	60	0.97914322
										0.91578475

Figure 8. Cosine vector similarity of same FoS papers

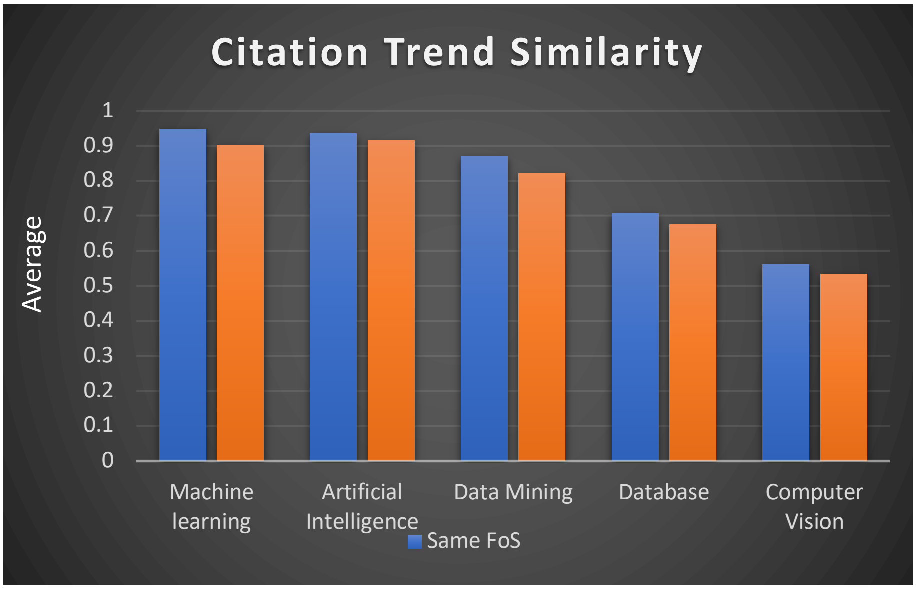

Figure 9. FoS Trend Similarity

The depicted Figure 9 presents the outcomes of analyzing citation trends within the same and different Fields of Study (FoS). Our experimentation encompasses six distinct FoS categories: "Artificial Intelligence," "Data Mining," "Machine Learning," "Computer Vision," and "Database". and The findings indicate that papers within the same FoS exhibit citation trends that closely align with the overall citation trend of that specific FoS, in contrast to papers from different FoS categories. This highlights the impact of the field of study on a paper's citation count, indicating that authors should take into account the current trends within a specific field when choosing their research area.

5. Conclusion and Recommendations

Assessing the impact of research trends holds paramount importance for research organizers, scholars, and policymakers. Given the varied prominence of Computer Science (CS) in contemporary groups and its active role in enhancing the efficiency and utility of structural processes, visualizing the landscape of scientific publications organized by scholars in this domain becomes crucial. This study offers a comprehensive overview of the widespread dissemination of CS publications by analyzing the significance of research trends in CS conference papers published between 1985 and 2017, as captured in the MAG.

Furthermore, the study examines the annual citation frequency for CS publications within MAG and discerns the citation trends among papers within the same field. It underscores the significance of adhering to research trends in the CS field and explores the impact of trends on research paper citations, particularly focusing on trends observed in CS conference papers. The findings highlight the substantial influence of the FoS on the citation count of a paper, emphasizing the importance for researchers to consider prevailing trends within their chosen research area. Additionally, the study reinforces that papers within the same Field of Study (FoS) exhibit similar citation trends, further underlining the importance of aligning research with relevant trends within the CS field.

5.1 Future Work

For Future research we will study what is the effect of following the popular trend on the careers of scientific authors. The effect of trend evaluation can be pragmatic to any science field to support cognize research trends and their evolvement in diverse fields. We will took an exploration on a set of active authors how trend effect authors ranking, who have incessant and persistent existence in the scientific literature across a duration of time.

Funding: This research received no external funding.

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

References

1. Pham MC, Klamma R, Jarke M. Development of computer science disciplines: a social network analysis approach. Social Network Analysis and Mining. 2011 Nov;1:321-40.
2. Hoonlor A, Szymanski BK, Zaki MJ. Trends in computer science research. Communications of the ACM. 2013 Oct 1;56(10):74-83.
3. Hicks D, Wouters P, Waltman L, De Rijcke S, Rafols I. Bibliometrics: the Leiden Manifesto for research metrics. Nature. 2015 Apr 23;520(7548):429-31.
4. Tseng YH, Lin YI, Lee YY, Hung WC, Lee CH. A comparison of methods for detecting hot topics. Scientometrics. 2009 Oct 1;81(1):73-90.
5. Zhang J, Yu Q, Zheng F, Long C, Lu Z, Duan Z. Comparing keywords plus of WOS and author keywords: A case study of patient adherence research. Journal of the association for information science and technology. 2016 Apr;67(4):967-72.
6. Garfield E. KeyWords Plus-ISI's breakthrough retrieval method. 1. Expanding your searching power on current-contents on diskette. Current contents. 1990 Aug 6;32:5-9.
7. Dong B, Xu G, Luo X, Cai Y, Gao W. A bibliometric analysis of solar power research from 1991 to 2010. Scientometrics. 2012 Dec 1;93(3):1101-17.
8. Wang X, Xu S, Peng L, Wang Z, Wang C, Zhang C, Wang X. Exploring scientists' working timetable: Do scientists often work overtime?. Journal of Informetrics. 2012 Oct 1;6(4):655-60.
9. Tan J, Fu HZ, Ho YS. A bibliometric analysis of research on proteomics in Science Citation Index Expanded. Scientometrics. 2014 Feb;98:1473-90.
10. Chen H, Jiang W, Yang Y, Man X, Tang M. A bibliometric analysis of waste management research during the period 1997–2014. Scientometrics. 2015 Nov;105:1005-18.
11. Yang Y, Wu M, Cui L. Integration of three visualization methods based on co-word analysis. Scientometrics. 2012 Feb 1;90(2):659-73.
12. Valderrama-Zurián JC, García-Zorita C, Marugán-Lázaro S, Sanz-Casado E. Comparison of MeSH terms and KeyWords Plus terms for more accurate classification in medical research fields. A case study in cannabis research. Information Processing & Management. 2021 Sep 1;58(5):102658.
13. Callon M, Courtial JP, Turner WA, Bauin S. From translations to problematic networks: An introduction to co-word analysis. Social science information. 1983 Mar;22(2):191-235.
14. Gupta BM, Bhattacharya S. Bibliometric approach towards mapping the dynamics of science and technology. DESIDOC Journal of Library & Information Technology. 2004;24(1).
15. Leydesdorff L, Schank T. Dynamic animations of journal maps: Indicators of structural changes and interdisciplinary developments. Journal of the American Society for Information Science and Technology. 2008 Sep;59(11):1810-8.
16. Leydesdroff L. Words and co-words as indicators of intellectual organization. Research policy. 1989 Aug 1;18(4):209-23.
17. Hui SC, Fong AC. Document retrieval from a citation database using conceptual clustering and co-word analysis. Online Information Review. 2004 Feb 1;28(1):22-32.
18. Brookshear JG, Smith D, Brylow D. Computer science: an overview.
19. Sinha A, Shen Z, Song Y, Ma H, Eide D, Hsu BJ, Wang K. An overview of microsoft academic service (mas) and applications. InProceedings of the 24th international conference on world wide web 2015 May 18 (pp. 243-246).
20. Hoonlor A, Szymanski BK, Zaki MJ. Trends in computer science research. Communications of the ACM. 2013 Oct 1;56(10):74-83.
21. Effendy S, Yap RH. Analysing trends in computer science research: A preliminary study using the microsoft academic graph. InProceedings of the 26th international conference on world wide web companion 2017 Apr 3 (pp. 1245-1250).
22. Gruhl D, Guha R, Liben-Nowell D, Tomkins A. Information diffusion through blogspace. InProceedings of the 13th international conference on World Wide Web 2004 May 17 (pp. 491-501).
23. Oka M, Abe H, Kato K. Extracting topics from weblogs through frequency segments. InProc. of the Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics 2006 May.
24. Morinaga S, Yamanishi K. Tracking dynamics of topic trends using a finite mixture model. InProceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining 2004 Aug 22 (pp. 811-816).
25. Cataldi M, Di Caro L, Schifanella C. Emerging topic detection on twitter based on temporal and social terms evaluation. InProceedings of the tenth international workshop on multimedia data mining 2010 Jul 25 (pp. 1-10).
26. Mathioudakis M, Koudas N. Twittermonitor: trend detection over the twitter stream. InProceedings of the 2010 ACM SIGMOD International Conference on Management of data 2010 Jun 6 (pp. 1155-1158).
27. Bolelli L, Ertekin Ş, Giles CL. Topic and trend detection in text collections using latent dirichlet allocation. InEuropean conference on information retrieval 2009 Apr 6 (pp. 776-780). Berlin, Heidelberg: Springer Berlin Heidelberg.
28. Decker SL, Aleman-Meza B, Cameron D, Arpinar IB. Detection of bursty and emerging trends towards identification of researchers at the early stage of trends.
29. Herrmannova D, Knoth P. An analysis of the microsoft academic graph. D-lib Magazine. 2016 Sep 15;22(9/10):1.
30. Erten C, Harding PJ, Kobourov SG, Wampler K, Yee G. Exploring the computing literature using temporal graph visualization. InVisualization and Data Analysis 2004 2004 Jun 4 (Vol. 5295, pp. 45-56). SPIE.
31. Lv PH, Wang GF, Wan Y, Liu J, Liu Q, Ma FC. Bibliometric trend analysis on global graphene research. Scientometrics. 2011 Aug 1;88(2):399-419.

32. Osborne F, Scavo G, Motta E. A hybrid semantic approach to building dynamic maps of research communities. InKnowledge Engineering and Knowledge Management: 19th International Conference, EKAW 2014, Linköping, Sweden, November 24-28, 2014. Proceedings 19 2014 (pp. 356-372). Springer International Publishing.
33. Sun X, Ding K, Lin Y. Mapping the evolution of scientific fields based on cross-field authors. Journal of Informetrics. 2016 Aug 1;10(3):750-61.
34. Osborne F, Motta E, Mulholland P. Exploring scholarly data with rexplore. InThe Semantic Web–ISWC 2013: 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21-25, 2013, Proceedings, Part I 12 2013 (pp. 460-477). Springer Berlin Heidelberg.
35. Salatino A. Early detection and forecasting of research trends.
36. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. Journal of machine Learning research. 2003;3(Jan):993-1022.
37. Blei D, Lafferty J. Correlated topic models. Advances in neural information processing systems. 2006 Dec;18:147.
38. Griffiths T, Jordan M, Tenenbaum J, Blei D. Hierarchical topic models and the nested Chinese restaurant process. Advances in neural information processing systems. 2003;16.
39. Chang J, Blei DM. Hierarchical relational models for document networks.