SCIENTIFIC OASIS

Decision Making Advances

Journal homepage: www.dma-journal.org
ISSN: 2956-2384

The Evolution of Sentiment Analysis Across Various Scientific Disciplines: A Comprehensive Review Based on the Bibliometric Technique

Abdullah Önden[1,*], Mohammed Alnour[2], Vladimir Simic[3], Dragan Pamucar[4]

1 Department of International Trade and Finance, Yalova University, 77200, Yalova, Turkey
2 Department of Economics, Erciyes University, Kayseri, Turkey
3 University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11010, Belgrade, Serbia
4 Department of Operations Research and Statistics, Faculty of Organizational Sciences, University of Belgrade, Belgrade, Serbia

ARTICLE INFO

Article history:
Received 3 December 2023
Received in revised form 25 April 2024
Accepted 2 June 2024
Published 20 June 2024

Keywords:
Sentiment Analysis; Bibliometric Analysis; VOSviewer; Web of Science; Systematic Review.

ABSTRACT

The unprecedented development of artificial intelligence has been associated with several challenges given its substantial influence on economic, social, and cultural structures. Particularly, natural language processing, text analysis, computational linguistics, and biometrics have been important topics of research among scholars in computer science (sentiment analysis), given the notable increase in social media marketing and e-commerce. In this respect, it becomes imperative to trace the evolution of sentiment analysis not only in computer science and information technology-related fields as existing literature does but also across other disciplines. Therefore, the current study is the first attempt to provide a comprehensive review for sentiment analysis using the bibliometric technique for 1668 research articles published in 2023 collected from Web of Science. In this study, we identify in percentage the contribution of sentiment analysis and common keywords across various research fields. Additionally, the most influential research and active collaborative authors are presented. Finally, the gaps and research topics, which received minimal attention regarding sentiment analysis, are suggested and elaborated on.

1. Introduction

Presently, artificial intelligence (AI) is by far the most prominent topic dominating world public opinion together with its important political, social, cultural, and demographic underpinnings [1,2]. Many analysts have widely agreed that the present dynamics and evolution of AI constitute a serious challenge to human societies [3]. AI has been used in almost every scientific discipline, e.g. education [4], medical sciences [5], entrepreneurship [6,7], social media [8], and economics [9,10].

[*] *Corresponding author.*
E-mail address: abdullah.onden@yalova.edu.tr

https://doi.org/10.31181/dma21202441

Historically, AI is older than normally known. It has roots in ancient Greece and spans several fields, including science and philosophy. But its modern incarnation owes a great deal to individuals such as Alan Turing and a seminal conference that took place at Dartmouth College in 1956, where John McCarthy officially introduced the term "artificial intelligence", defining it as "the science and engineering of creating intelligent machines". Since then, AI has experienced numerous ups and downs since its early inception in the 1950s. These periods are commonly referred to as AI "summer and winters". Starting around 2010, it could be argued that AI entered a "summer period" principally motivated by significant advances in computer processing power and the accessibility of large datasets [11].

The revival of interest in AI research can be attributed to three major advances: (a) the development of a more sophisticated class of algorithms; (b) the introduction of affordable graphics processors capable of performing extensive calculations in milliseconds; and (c) the availability of vast, accurately annotated databases, allowing the development of more advanced intelligent systems [12]. Computer sciences and related fields have shown a substantial increase in the utilization of different AI classes. Predicting sentiments and emotions from people’s texts is an essential concern in cognitive computing at the current time. Over recent years “sentiment analysis” has evolved substantially not only in the computer sciences research field but also proliferated into various other social, natural, and health research areas [13].

Sentiment analysis, also known as opinion mining or emotion AI, is the systematic identification, extraction, quantification, and study of affective states and subjective information using natural language processing, text analysis, computational linguistics, and biometrics [14]. Sentiment analysis, also known as opinion mining, is the process of the evaluation of one's opinion, sentiment, and attitude toward an entity, as stated in written works [15]. Because of the numerous studies and research findings that are continuously available in academic settings, sentiment analysis has consequently emerged as a vibrant area within information processing [13]. Therefore, examining sentiment analysis's trends and the current state is important and required given the diversity of scientific literature on the subject, especially about the main research questions that academics are interested in. In this vein, the key motives that lie behind this review are to provide a more holistic view of the research with “sentiment analysis”, identify the uncovered topics, and suggest future research agenda.

Theoretically, an extensive line of literature provided valuable insights on sentiment analysis-related studies, e.g. Brauwers and Frasincar [16] conducted a comprehensive review following the persistent increase in the demand for automatic sentiment analysis algorithms. Motivated by the recent COVID-19 outbreak and its implication on news and the stock market, Costola et al. [17] provided sentiment analysis through a financial market-adapted BERT model that enables recognizing the context of each word in each item. Additionally, Mao et al. [18] carried out a thorough empirical investigation on prompt-based sentiment analysis and emotion detection in order to examine PLM biases in the direction of affective computing. However, most of the survey and review studies concentrated greatly on sentiment analysis within computer sciences and related fields such as emotion classification, pre-trained language models, conversational interfaces, accuracy, meta-based self-training, semi-supervised multi-modal, and ignored the evolution of sentiment analysis across other disciplines.

Therefore, the main objective of this study is to provide a comprehensive review of the evolution of sentiment analysis across different research fields. In doing so, this study employs a bibliometric technique for 1668 high-quality research articles published in 2023 extracted from the Web of Science (WoS). Following the WoS database, the research with sentiment analysis reached a record in 2023. Thus, focusing on 2023 can bring important insights and guide future related research.

Bibliometric analysis is a valid and dependable method for assessing scientific outputs, especially in the age of "big data". This technique provides several contributions to the existing literature such as (i) identifying common research keywords across various research domains; (ii) identifying the most influential research; and (iii) identifying the most active researchers in the relevant fields. Furthermore, this analysis also allows us to identify the future research areas that are not widely addressed. These advantages are the core target of the present review paper to explain the development of sentiment analysis and to explore some future insights. We believe that the evidence will greatly enrich literature and offer valuable information for researchers in pertinent fields.

The remaining part of this research is arranged as follows: the next part outlines the related literature. The third part elaborates on methodology. Meanwhile, the final parts portray the results followed by a conclusion.

2. Literature Review

Sentiment analysis can be defined as a branch of text classification, defined as the process of extracting emotion terms such as mood or emotional tone in text data and determining the semantic orientation of the opinion [19]. With this method, machine learning and natural language processing techniques are used to examine text data and classify it into emotional categories such as positive, negative, or neutral [20]. In research conducted with sentiment analysis, for instance, Onden and Calli [21] developed a framework for scrutinizing user interactions on social media through natural language processing (NLP). These algorithms are used to make sense of emotional content by processing large amounts of data and to detect hidden opinions in documents. For this reason, sentiment analysis has become an important research field to understand and evaluate clear and ambiguous elements in texts through the combination of text mining and natural language processing techniques [13]. There are many studies in the literature about the use of sentiment analysis in various fields. These studies evaluate the current situation and explain the development trends in this field.

In their study, Nandwani and Verma [22] focused on understanding the levels of emotion analysis, different emotion models, and the process of identifying emotions from texts. On the other hand, they examined this field from a broader perspective by addressing the difficulties encountered during emotion and emotion analysis. Birjali et al. [23] evaluated sentiment analysis and the approaches, challenges, and future trends in this field in detail in their research. In this context, in the research, the usage areas and general process of sentiment analysis were touched upon, various approaches were examined and compared, and the advantages and disadvantages of sentiment analysis were discussed. The challenges and future trends of sentiment analysis are also discussed in this research. Wankhade et al. [24] examined how sentiment analysis is structured to extract subjective information from texts, its applications, and the advantages and disadvantages of different approaches used in this field. They also discussed the challenges and future directions that may be encountered in the sentiment analysis process.

Machine learning techniques are often preferred by researchers to analyze different emotional expressions on large data sets. These techniques are used to understand, classify, and analyze the emotional content of texts. In fields such as text mining and natural language processing, various machine learning techniques focus on recognizing and interpreting emotional expressions [25]. Rodrigues et al. [26] tried to develop a system that can detect whether tweets are spam or real and evaluate the sentiment of the tweet. For this purpose, the features obtained by pre-processing the tweets were evaluated using various classifiers, and different deep-learning methods were used for sentiment analysis.

Singh et al. [27] used four machine learning classifiers consisting of Naive Bayes, J48, BFTree, and OneR for sentiment analysis optimization. The effectiveness of these four classification techniques in the research data set was examined and compared. Manek et al. [28] proposed a Gini index-based feature selection method together with the support vector machine (SVM) classifier to perform sentiment classification. Huq et al. [29] aimed to develop a functional classifier that can accurately and automatically classify the sentiment tags of an unknown tweet using two different methods. Additionally, they evaluated the performance of these methods, consisting of K-nearest neighbor (KNN)-based sentiment classification algorithm (SCA) and SVM, on real tweets sent by users.

Bibliometric analysis is an extremely effective method widely used to explore and examine large amounts of scientific data. This analysis not only reveals the evolutionary nuances in a particular discipline but also enables the discovery of new and important areas of development in that discipline. In this way, it helps understand the development of the scientific field over time, while also revealing the potential for discoveries [30]. The increase in information processing power and the diversity of accessible analytical tools allowed bibliometric analysis applications to develop rapidly in recent years and be widely used to show the evolution patterns of scientific literature in different disciplines [31]. For example, Sharifi et al. [32] evaluated the data on 105,015 documents published between 1956 and 2022 by using the bibliometric analysis method and determined the main themes of the field of urbanism and examined how they evolved. Fan et al. [33] carried out both bibliometric and discourse analyses of scientific research on large language models. In this context, the research reveals the current status, development, and application areas of large language models in scientific literature. Farrukh et al. [34] analyzed the trends for pro-environmental behavior research based on performance and science mapping using bibliographic data and evaluated the directions of future research in this field.

Systematic review refers to the detailed research, organization, and evaluation of the existing literature using systematic methods to find an answer to a specific research question. With this method, all available documents related to the research topic are compiled and analyzed [35]. Systematic review requires a narrow scope of study. This suggests that it tends to include fewer articles for review. For this reason, a systematic review is considered a more suitable re-view method for limited or niche research areas [30]. The combined use of bibliometric analysis and systematic review methods allows scientific activity and trends in a particular research field to be addressed more comprehensively by combining the information it provides with systematic depth with the quantitative data of bibliometric analysis. This offers researchers both a comprehensive and detailed perspective on the subject.

Yıldırım et al. [36] conducted a concept-oriented and systematic review of academic studies published in business literature between 2000 and 2022 to reveal a general view of marketing ethics literature. Ezugwu et al. [37] systematically examined the developments and trends in meta-heuristic clustering methods in the period between 1990 and 2019, presented a taxonomic overview, and carried out a bibliometric analysis. By combining systematic review and bibliometric analysis on financial literacy, Goyal and Kumar [38] aimed to provide both quantitative and qualitative information on this subject. Within the scope of this study, they examined 502 articles published between 2000 and 2019 and used methods such as content analysis, citation network, and publication trends to explain the intellectual structure of the field and identify gaps.

3. Literature Review

The integration of bibliometric and citation analysis provides a strong method for examining the characteristics and frameworks of previous academic literature in a variety of subject areas. This methodology is also useful for identifying the dominant schools of thought in the fields of study. The bibliometric analysis uses a quantitative research methodology and an objective philosophical framework to examine written materials including books, journals, and internet resources [12]. Citation and co-citation analyses mostly focus on detecting new themes in each field of study and evaluating the impact of different publications and schools of thinking. Previous studies have highlighted the ability to assess an academic field's progress and substance in ways other than just tallying and combining citations. This evaluation considers the achievements of specific publications and writers who have worked together to provide a collaborative contribution to the scientific community. Co-citation in a bibliographic analysis helps analyze the patterns and details of advances found in the literature, supporting the evaluation, organization, and articulation of earlier studies on a certain topic [39].

In general, the newly developed program VOSviewer is used to conduct bibliometric evaluation given its strong construction and visualization capabilities characteristics. It creates, visualizes, and analyzes these networks using a variety of components. These components may include words, writers, or works that are representative of the subjects of interest. Links can create relationships between any two items; they can be based on the co-occurrence of a term or concept, be found in a publication's references, or be found in a researcher's co-authorship links. Any link's strength can be expressed as a positive numerical value that goes down for weakened connections and up for stronger ones. The number of publications that are co-authored by two or more scholars, or the number of publications that are referenced in common by two or more publications, can be shown by the strength of each particular connection. Building different kinds of networks is made feasible by the connections between pieces and their links [12, 39].

Furthermore, the attendant literature can be followed, counted, and analyzed using bibliographic analysis. It includes a list of the researchers' published works, the refereed journals, the technique used, and concluding remarks. Utilizing metadata, any area of research can be condensed. The use of bibliometric methodologies, which necessitate a substantial amount of bibliographic data, has allowed for the analysis of a wide range of topics, journals, and nations [4]. This part briefly highlights the outcomes from the bibliometric analysis for 1668 sentiment research items using data collected from the WoS database. The survey of literature in the WoS database included all the fields using "sentiment analysis" as the main keyword. Initially, the analytical framework is shown in Figure 1.

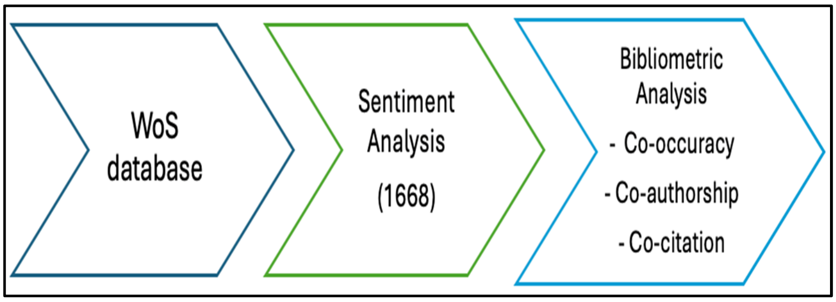

Fig. 1. Analysis flowchart.

The search for research on sentiment analysis through WoS resulted in around 1668 research articles with 1235 cited articles, 1090 without self-citations, and 1447 articles with times cited Figure 2 highlights graphically some insights about the existing relevant literature. Review articles, book chapters, editorial materials, and proceeding papers were excluded.

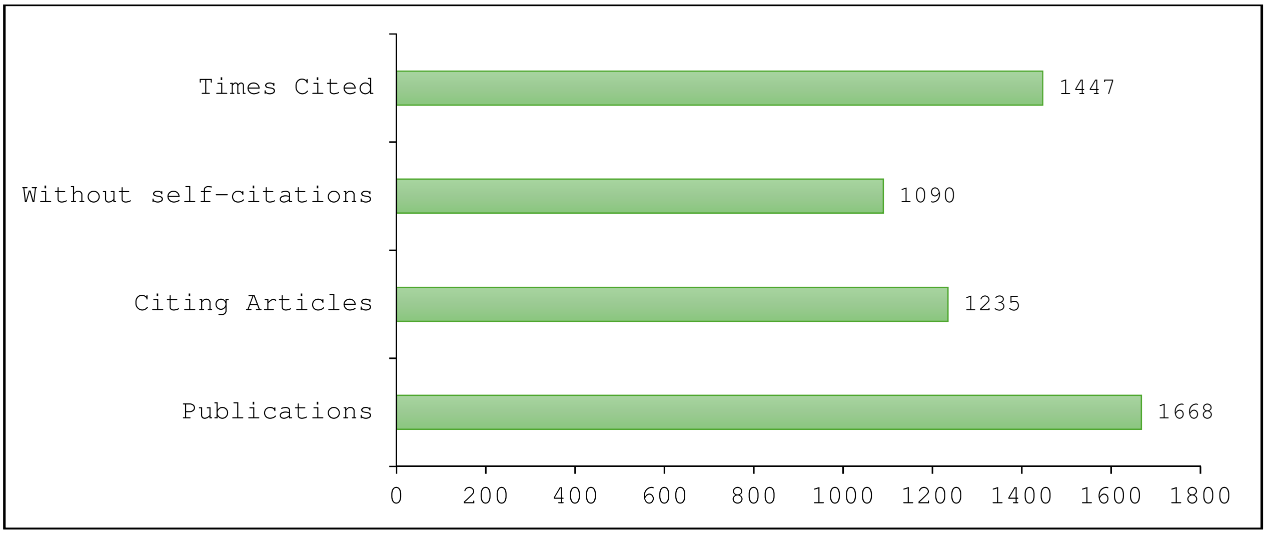

Fig. 2. Sentiment-related research characteristics.

4. Results and Discussion

First, the distribution of sentiment analysis in the scientific research fields was evaluated. Historically, sentiment analysis has increased substantially over the last three years, especially after 2022. Figure 3 outlines the development of research on sentiment analysis from a historical viewpoint.

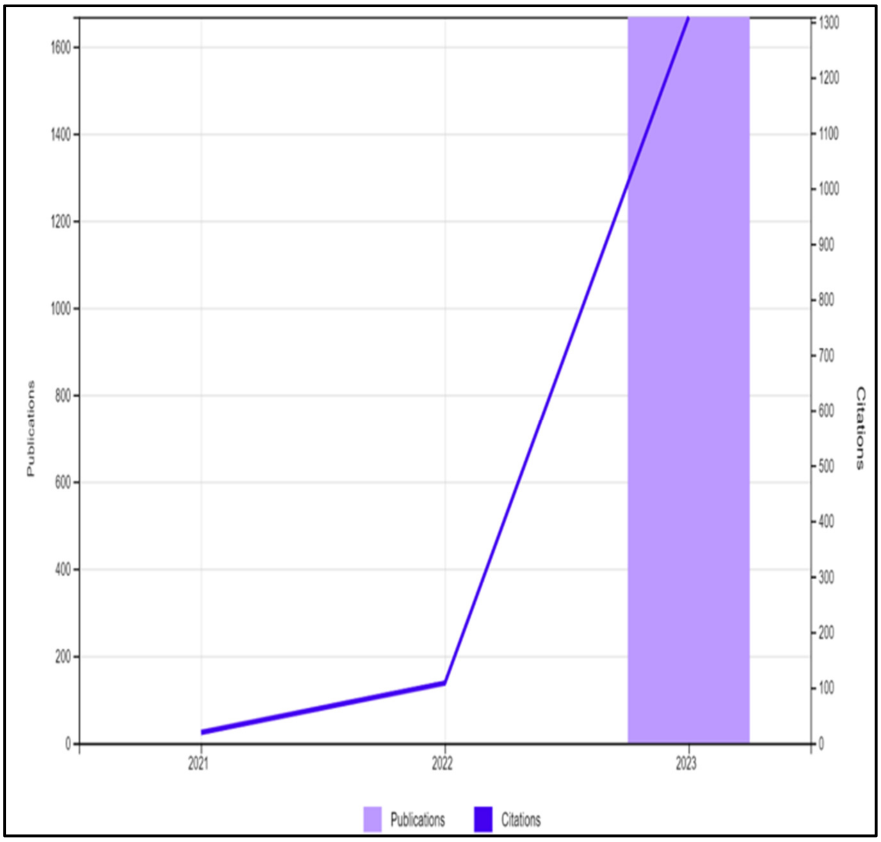

Fig. 3. The development of sentiment analysis in the research.

Table 1 provides sentiment analysis across different fields, with a wide range of scientific literature concerning sentiment analysis, computer science, information systems, and related fields. AI, electrical electronics, telecommunications, software engineering, information systems, and cybernetics are the disciplines utilized the sentiment analysis.

Table 1
Fields of categories out of 1,668 publications on sentiment analysis

Field	Record count	% of 1,668
Computer science information systems	439	26.32%
Computer science artificial intelligence	353	21.16%
Engineering electrical electronic	266	15.95%
Computer science theory methods	209	12.53%
Telecommunications	104	6.24%
Engineering multidisciplinary	100	6.00%
Computer science software engineering	92	5.52%
Physics applied	82	4.92%
Computer science interdisciplinary applications	72	4.32%
Multidisciplinary sciences	68	4.08%
Materials science multidisciplinary	67	4.02%
Business	62	3.72%
Information science library science	59	3.54%
Management	50	3.00%
Health care sciences services	49	2.94%
Chemistry multidisciplinary	48	2.88%
Computer science cybernetics	48	2.88%
Environmental sciences	47	2.82%
Automation control systems	36	2.16%
Medical informatics	36	2.16%
Environmental studies	35	2.10%
Operations research management science	35	2.10%
Education educational research	33	1.98%
Communication	32	1.92%
Computer science hardware architecture	32	1.92%
Public environmental occupational health	32	1.92%
Hospitality leisure sport tourism	31	1.86%
Green sustainable science technology	30	1.80%
Economics	28	1.68%
Social sciences interdisciplinary	28	1.68%
Business finance	25	1.50%
Health policy services	19	1.14%
Neurosciences	17	1.02%
Instruments instrumentation	16	0.96%
Political science	16	0.96%
Chemistry analytical	15	0.90%
Engineering industrial	15	0.90%
Mathematics interdisciplinary applications	15	0.90%
Mathematics	14	0.84%
Psychology multidisciplinary	13	0.78%
Linguistics	11	0.66%
Water resources	11	0.66%
Humanities multidisciplinary	10	0.60%
Psychology experimental	10	0.60%
Meteorology atmospheric sciences	9	0.54%
Geography	8	0.48%
Immunology	8	0.48%

Table 1 (*continued*)

Field	Record count	% of 1,668
International relations	8	0.48%
Medicine research experimental	8	0.48%
Transportation	8	0.48%
Area studies	7	0.42%
Geosciences multidisciplinary	7	0.42%
Physics multidisciplinary	7	0.42%
Psychology social	7	0.42%
Social sciences mathematical methods	7	0.42%
Sociology	7	0.42%
Transportation science technology	7	0.42%
Acoustics	6	0.36%
Education scientific disciplines	6	0.36%
Engineering civil	6	0.36%
Engineering manufacturing	6	0.36%
Ergonomics	6	0.36%
Language linguistics	6	0.36%
Mathematics applied	6	0.36%
Psychiatry	6	0.36%
Public administration	6	0.36%
Regional urban planning	6	0.36%
Statistics probability	6	0.36%
Biodiversity conservation	5	0.30%
Biotechnology applied microbiology	5	0.30%
Clinical neurology	5	0.30%
Ecology	5	0.30%
Engineering biomedical	5	0.30%
Food science technology	5	0.30%
Medicine general internal	5	0.30%
Construction building technology	4	0.24%
Engineering chemical	4	0.24%
Engineering environmental	4	0.24%
Mathematical computational biology	4	0.24%
Oncology	4	0.24%
Psychology applied	4	0.24%
Social issues	4	0.24%
Surgery	4	0.24%
Agronomy	3	0.18%
Demography	3	0.18%
Dentistry oral surgery medicine	3	0.18%
Energy fuels	3	0.18%
Forestry	3	0.18%
Oncology	4	0.24%
Obstetrics gynecology	3	0.18%
Otorhinolaryngology	3	0.18%
Plant sciences	3	0.18%
Psychology mathematical	3	0.18%
Radiology nuclear medicine medical imaging	3	0.18%
Remote sensing	3	0.18%
Substance abuse	3	0.18%
Agriculture multidisciplinary	2	0.12%
Biology	2	0.12%
Criminology penology	2	0.12%
Cultural studies	2	0.12%
Geography physical	2	0.12%

Table 1 (*continued*)

Field	Record count	% of 1,668
History	2	0.12%
Imaging science photographic technology	2	0.12%
Law	2	0.12%
Microbiology	2	0.12%
Nuclear science technology	2	0.12%
Optics	2	0.12%
Pediatrics	2	0.12%
Psychology clinical	2	0.12%
Agricultural economics policy	1	0.06%
Biophysics	1	0.06%
Chemistry applied	1	0.06%
Development studies	1	0.06%
Engineering aerospace	1	0.06%
Engineering mechanical	1	0.06%
Film radio television	1	0.06%
Fisheries	1	0.06%
Genetics heredity	1	0.06%
History of social sciences	1	0.06%
Industrial relations labor	1	0.06%
Literature	1	0.06%
Literature romance	1	0.06%
Marine freshwater biology	1	0.06%
Medical ethics	1	0.06%
Medieval renaissance studies	1	0.06%
Nursing	1	0.06%
Nutrition dietetics	1	0.06%
Orthopedics	1	0.06%
Pharmacology pharmacy	1	0.06%
Physics mathematical	1	0.06%
Physics particles fields	1	0.06%
Psychology educational	1	0.06%
Quantum science technology	1	0.06%
Rehabilitation	1	0.06%
Religion	1	0.06%
Reproductive biology	1	0.06%
Robotics	1	0.06%
Social sciences biomedical	1	0.06%
Social work	1	0.06%
Sport sciences	1	0.06%
Toxicology	1	0.06%
Urban studies	1	0.06%
Veterinary sciences	1	0.06%
Women studies	1	0.06%

According to Chen and Xie [13], predicting sentiments and emotions from people's texts was an essential concern in cognitive computing fields of research. These above disciplines are also generally consistent with the most used keywords in the research on sentiment analysis as highlighted in Figure 4.

Determining keywords across different disciplines is one of the most useful ways to track the evolution of a specific method [12]. Figure 4 portrays words that appear in the same article at the same time. It is worth mentioning that the abstract only considers the keywords specified by the author. In addition, the figure also presents the most relevant keywords highlighted, as well as the nodes linking them. Weight increases with the size of the term and node; the narrower the gap between nodes, the stronger the association between them. The author obtained 28 keywords by using the full counting method and a co-occurrence analysis type of phrase with at least 20 co-occurrences.

As can be seen clearly, the computer sciences and related fields are dominating the research on sentiment analysis. Precisely, the most apparent keywords include "deep learning, natural language processing, social media, Twitter, machine learning, classification, COVID-19, model, Istm, BERT, impact, text mining, attention, topic modeling, opinion mining, aspect-based sentiment analysis, AI, information task analysis, feature extraction, reviews, attention mechanism, NLP, sentiment, online reviews, multimodal sentiment analysis".

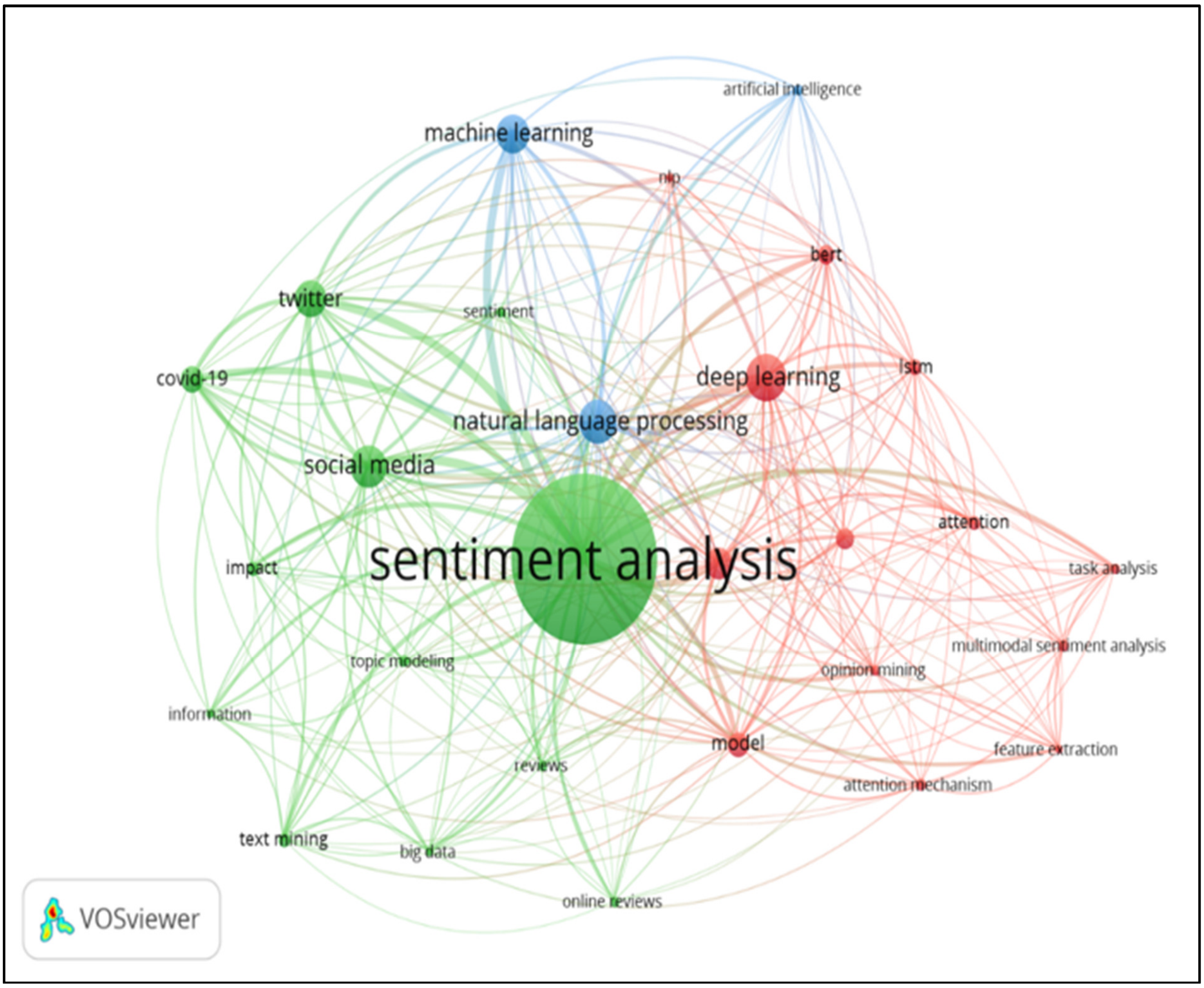

Fig. 4. Co-occurrence (keywords) in the research on sentiment analysis.

The visualization and density in Figure 5 and Figure 6 show closely how the research develops over time using the keyword's popularity. The more yellow the cluster, the more recent the research. Most of the keywords that are highlighted in yellow show how the evolution of research related to "sentiment analysis". The density of associated terms and the quantity of nodes near each other. These findings can be utilized as a foundation for future research on sentiment analysis when combined with other areas of science. A saturation level represented by a yellow node indicates the prevalence of frequently occurring phrases.

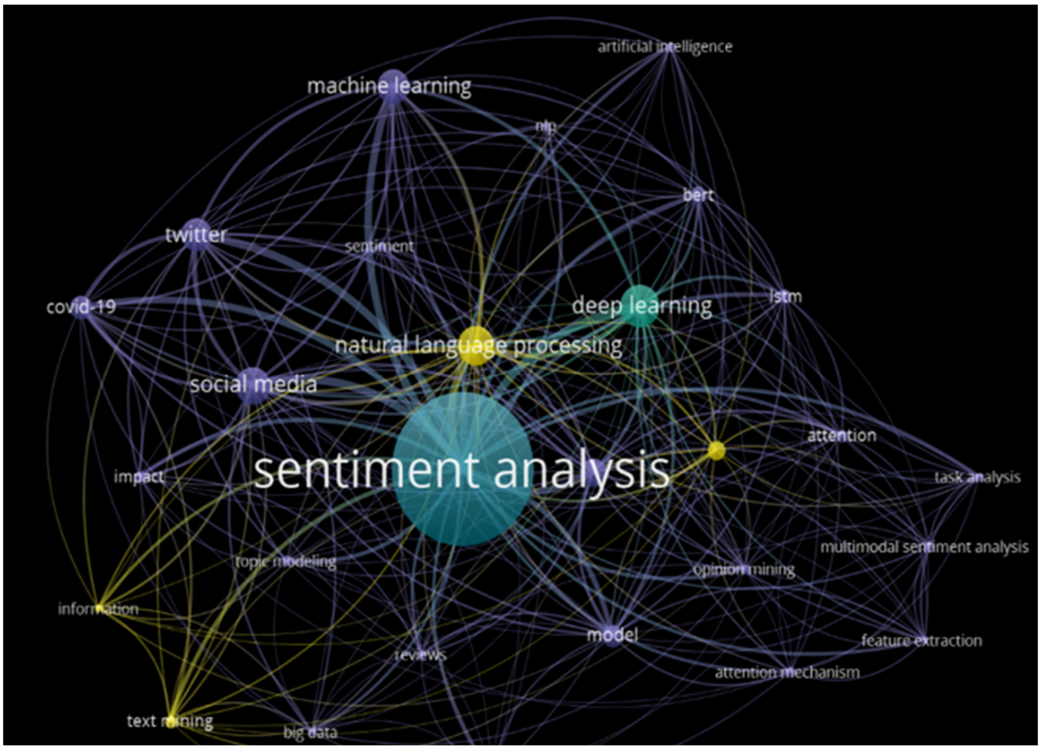

Fig. 5. Overlay visualization co-occurrence (keywords).

In the case where the keyword "sentiment analysis" is most found, the yellow node signifies the region with the most extensive research. Different topics are identified, such as "deep learning, multimodal sentiment analysis, feature extraction, opinion mining, big data, and text mining". These topics could be the focus of future "sentiment analysis" research.

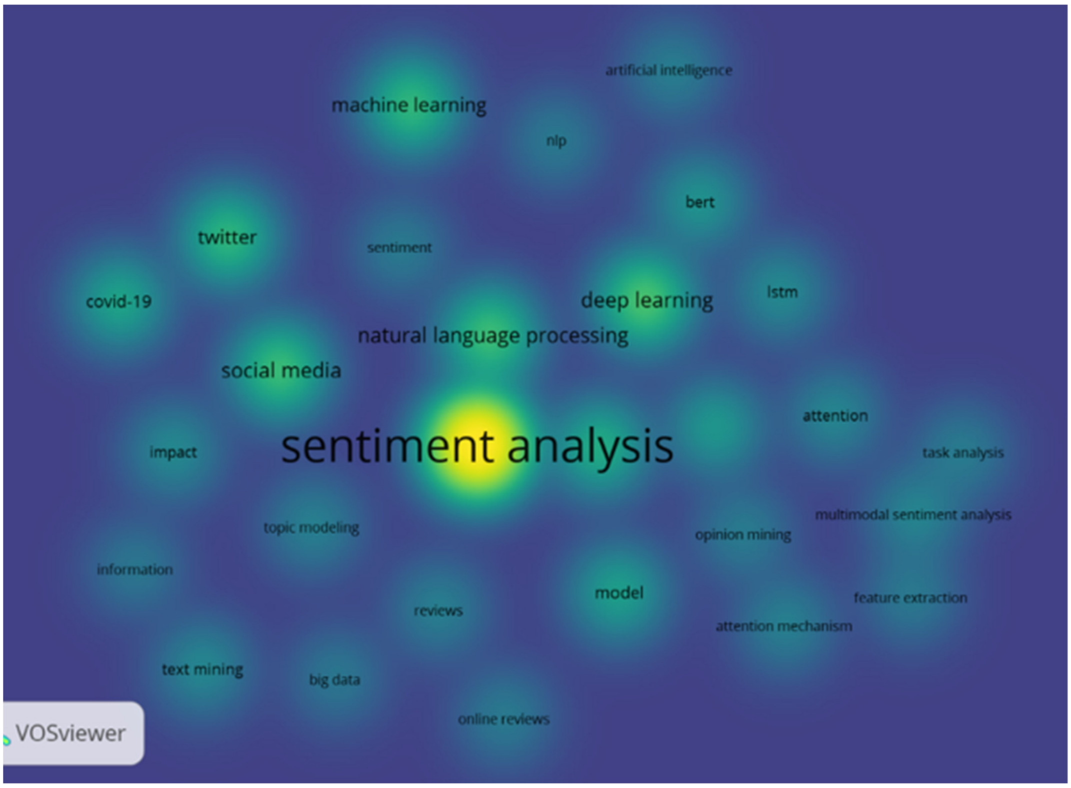

Fig. 6. Density visualization co-occurrence (keywords).

Another important goal for the current study is to determine the most impactful research along with the scholars that greatly contributed to the fields of research using "sentiment analysis". Table 2 in brief summarizes the most influential research on sentiment analysis. With a minimum number of three articles, only 42 authors met the threshold of 3381 authors.

Table 2
Most influential research on sentiment analysis in 2023

Author(s)	Title	Total citations
Liu et al. [40]	Emotion classification for short texts: an improved multi-label method	60
Zhang et al. [41]	A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges	29
Poria et al. [42]	Beneath the Tip of the Iceberg: Current Challenges and New Directions in Sentiment Analysis Research	27
Mao et al. [18]	The Biases of Pre-Trained Language Models: An Empirical Study on Prompt-Based Sentiment Analysis and Emotion Detection	22
Hassani and Silva [43]	The Role of ChatGPT in Data Science: How AI-Assisted Conversational Interfaces Are Revolutionizing the Field	21
Hartmann et al. [44]	More than a Feeling: Accuracy and Application of Sentiment Analysis	20
Brauwers and Frasincar [16]	A Survey on Aspect-Based Sentiment Classification	18
Hu et al. [45]	Vaccine supply chain management: An intelligent system utilizing blockchain, IoT and machine learning	16
Bouschery et al. [46]	Augmenting human innovation teams with artificial intelligence: Exploring transformer-based language models	16
Zhu et al. [47]	Multimodal Sentiment Analysis with Image-Text Interaction Network	15
Jin et al. [48]	Back to common sense: Oxford dictionary descriptive knowledge augmentation for aspect-based sentiment analysis	12
Bello et al. [49]	A BERT Framework to Sentiment Analysis of Tweets	12
He et al. [50]	Meta-Based Self-Training and Re-Weighting for Aspect-Based Sentiment Analysis	11
Liu et al. [51]	Public perceptions of environmental, social, and governance (ESG) based on social media data: Evidence from China	11
Lian et al. [52]	SMIN: Semi-Supervised Multi-Modal Interaction Network for Conversational Emotion Recognition	10
Thakur [53]	Sentiment Analysis and Text Analysis of the Public Discourse on Twitter about COVID-19 and MPox	10
Alqarni and Rahman [54]	Arabic Tweets-Based Sentiment Analysis to Investigate the Impact of COVID-19 in KSA: A Deep Learning Approach	9
Wang et al. [55]	Towards a Robust Deep Neural Network Against Adversarial Texts: A Survey	9
Costola et al. [17]	Machine learning sentiment analysis, COVID-19 news and stock market reactions	9
Li et al. [56]	A Study on Public Perceptions of Carbon Neutrality in China: has the Idea of ESG Been Encompassed?	9

Figure 7 shows the most active authors in the research fields with sentiment analysis. These can be listed as Ambria Erik, Ashraf Imran, Rustam Furqan, Hassan Abdulkhaleq, Motwakel Abdel-wahed, Mujahid Muhammad, Alshahrani Hala, Mao Rul, Yaseen Ishfaq, An Jieyu, Aravapalli Rama Satish, Das Ringki, Feng Jun, Jain Goonjan, Krosuri Lakshmi Revathi, Li Wei, Praveen, Punetha Neha, Ribeiro-Soriano Domingo, Saura Ose, Sharaff Aakanksha, Singh Thoudam Doren, Srinivasarao Ulligaddala, Sun Xla, Vajrobol Vajratiya, Zainon Wan, Mohd Nazmee Wan, Conversano Claudio, Romano Maurızio, Wang Zhaoxia, Cal Qianhua, Xue Yun, Chen Chen, El-Ansarı Anas, Herrera-Viedma Enrique, Li Zuhe, Iu Ning, Nayyar Anand, Sufi Fahim, Thakur Nirmalya, Wang Ke, Zhang Jian, and Zhang Jing.

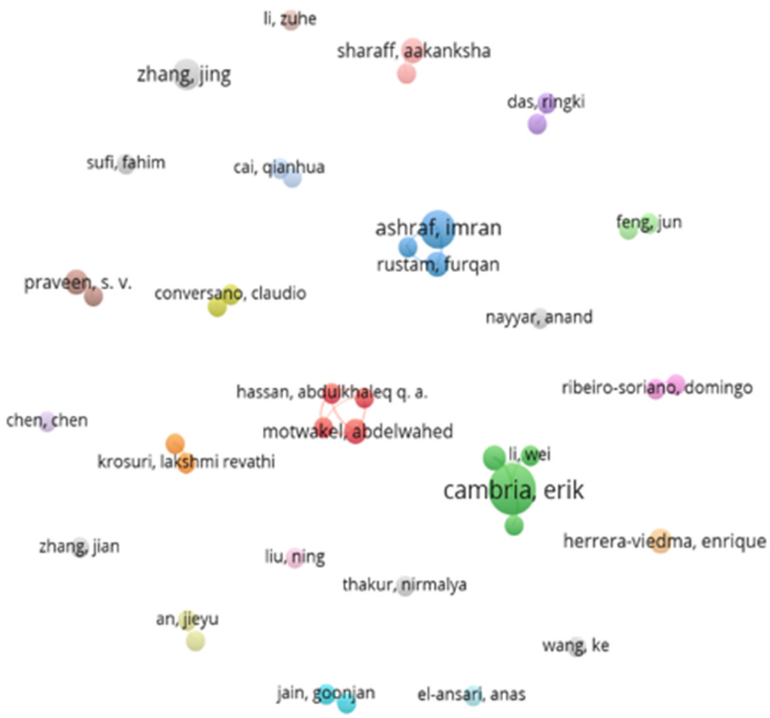

Fig. 7. Most active authors.

5. Research Gaps

The need for automatic sentiment analysis algorithms is only going to increase due to the proliferation of reviews and other texts on the Web that carry out the sentiment. Several research works attempted to evaluate the development of sentiment analysis. Yet, the computer sciences and its related fields have dominated the research fields on sentiment analysis and remain silent on the proliferation of sentiment research across other disciplines. By an in-depth search, one can claim that there is a great need to evaluate the evolution of sentiment research more holistically by identifying some common keywords, the most influencing relevant studies, effective organizations, and active researchers. More importantly, it highlights the research areas that receive little attention within sentiment analysis. This study, therefore, is a pioneering attempt to respond to the need for a comprehensive review of sentiment research.

6. Conclusions

Sentiment analysis is a rapidly evolving area of computer science research. With such many studies on innovative sentiment analysis available, it is worthwhile to present a review to fully comprehend the research on sentiment analysis. With the unprecedented level of diffusion and competition for developing more advanced and sophisticated classes of AI, huge attention has been given to the research on cognitive computing which has been raising many significant concerns in academia and policymaking. This necessitates a clear direction that specifies how it can contribute, as well as identifying the key research themes being covered and future research agenda.

In this respect, the current study aimed to contribute to the ongoing discussion by reviewing the attendant literature to identify the areas of interest pertinent to "sentiment analysis" and exploring some insight for future related research. In light of that, the current research conducted a comprehensive biblio-graphic review of 1668 articles generated from the WoS database. The analysis was conducted based on three dimensions, namely, co-occurrence and research distribution, co-authorship, and research impact.

We found that the most influenced research on sentiment analysis centered on emotion classification, pre-trained language models, conversational interfaces, accuracy, vaccine supply chain management, augmenting human innovation, multimodal sentiment analysis, knowledge augmentation, meta-based self-training, public perceptions, semi-supervised multi-modal, public discourse, deep learning approach, and machine learning. It is also worth mentioning that the most frequent keywords related to sentiment analysis research are deep learning, natural language processing, social media, Twitter, machine learning, classification, COVID-19, model, lstm, BERT, impact, text mining, attention, topic modeling, and opinion mining. We also found that out of 3381 authors contributed to the research field on sentiment analysis. Around 42 authors published over three articles with 10 as the highest number being recorded per author.

It is recommended that further research using sentiment analysis be conducted on deep learning, multimodal sentiment analysis, feature extraction, opinion mining, big data, and text mining since they have received minimal attention.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Hagerty, A., & Rubinov, I. (2019). Global AI ethics: a review of the social impacts and ethical implications of artificial intelligence. arXiv preprint arXiv:1907.07892. https://doi.org/10.48550/arXiv.1907.07892.

[2] Gallego, A., & Kurer, T. (2022). Automation, digitalization, and artificial intelligence in the workplace: implications for political behavior. *Annual Review of Political Science, 25*, 463–484. https://doi.org/10.1146/annurev-polisci-051120-104535.

[3] Önden, A., & Alnour, M. (2023). ChatGPT and OpenAI: A Comprehensive Bibliometric Review. *Journal of Soft Computing and Decision Analytics, 1*(1), 254-264. https://doi.org/10.31181/jscda11202324.

[4] Pradana, M., Elisa, H. P., & Syarifuddin, S. (2023). Discussing ChatGPT in education: A literature review and bibliometric analysis. *Cogent Education, 10*(2), 2243134. https://doi.org/10.1080/2331186X.2023.2243134.

[5] Dave, T., Athaluri, S. A., & Singh, S. (2023). ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Frontiers in Artificial Intelligence, 6*, 1169595. https://doi.org/10.3389/frai.2023.1169595.

[6] Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023). ChatGPT has aced the test of understanding in college economics: Now what?. *The American Economist, 68*(2), 233-245. https://doi.org/10.1177/05694345231169654.

[7] Short, C. E., & Short, J. C. (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights, 19*, e00388. https://doi.org/10.1016/j.jbvi.2023.e00388.

[8] Biswas, S. (2023). The Function of chat GPT in Social Media: According to chat GPT. Available at SSRN 4405389. https://doi.org/10.2139/ssrn.4405389.

[9] Ali, H., & Aysan, A. F. (2023). What will ChatGPT revolutionize in financial industry?. Available at SSRN 4403372. https://doi.org/10.2139/ssrn.4403372.

[10] Zarifhonarvar, A. (2024). Economics of chatgpt: A labor market view on the occupational impact of artificial intelligence. *Journal of Electronic Business & Digital Economics, 3*(2), 100-116. https://doi.org/10.1108/JEBDE-10-2023-0021.

[11] Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management, 60*, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383.

[12] Önden, A. (2023). Artificial Intelligence in Management Information System: Bibliographical Review. In: *International Research in Social, Human and Administrative Sciences XVIII, Istanbul: Eğitim Yayınevi*, pp. 41–56.

[13] Chen, X., & Xie, H. (2020). A structural topic modeling-based bibliometric study of sentiment analysis literature. *Cognitive Computation, 12*, 1097-1129. https://doi.org/10.1007/s12559-020-09745-1.

[14] Qazi, A., Raj, R. G., Hardaker, G., & Standing, C. (2017). A systematic literature review on opinion types and sentiment analysis techniques: Tasks and challenges. *Internet Research, 27*(3), 608-630. https://doi.org/10.1108/IntR-04-2016-0086.

[15] Tirea, M. (2013). Traders' Behavior Effect on Stock Price Evolution. In: *15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*, pp. 273–280. https://doi.org/10.1109/SYNASC.2013.43.

[16] Brauwers, G., & Frasincar, F. (2022). A survey on aspect-based sentiment classification. *ACM Computing Surveys, 55*(4), 1-37. https://doi.org/10.1145/3503044.

[17] Costola, M., Hinz, O., Nofer, M., & Pelizzon, L. (2023). Machine learning sentiment analysis, COVID-19 news and stock market reactions. *Research in International Business and Finance, 64*, 101881. https://doi.org/10.1016/j.ribaf.2023.101881.

[18] Mao, R., Liu, Q., He, K., Li, W., & Cambria, E. (2022). The biases of pre-trained language models: An empirical study on prompt-based sentiment analysis and emotion detection. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2022.3204972.

[19] Tubishat, M., Idris, N., & Abushariah, M. A. (2018). Implicit aspect extraction in sentiment analysis: Review, taxonomy, oppportunities, and open challenges. *Information Processing & Management, 54*(4), 545-563. https://doi.org/10.1016/j.ipm.2018.03.008.

[20] Mohammad, S. M. (2016). Sentiment analysis: Detecting valence, emotions, and other affectual states from text. In: *Emotion measurement* (pp. 201-237). Woodhead Publishing.

[21] Onden, A., & Kiygi-Calli, M. (2018). The effects of brands'posts on social media: conceptual framework and application method. *Research Journal of Business and Management, 5*(3), 238-250.

[22] Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining, 11*(1), 81. https://doi.org/10.1007/s13278-021-00776-6.

[23] Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems, 226*, 107134. https://doi.org/10.1016/j.knosys.2021.107134.

[24] Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review, 55*(7), 5731-5780. https://doi.org/10.1007/s10462-022-10144-1.

[25] Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access, 8*, 23522-23530. https://doi.org/10.1109/ACCESS.2020.2969854.

[26] Rodrigues, A. P., Fernandes, R., Shetty, A., K, A., Lakshmanna, K., & Shafi, R. M. (2022). [Retracted] Real-Time Twitter Spam Detection and Sentiment Analysis using Machine Learning and Deep Learning Techniques. *Computational Intelligence and Neuroscience, 2022*(1), 5211949. https://doi.org/10.1155/2022/5211949.

[27] Singh, J., Singh, G., & Singh, R. (2017). Optimization of sentiment analysis using machine learning classifiers. *Human-centric Computing and information Sciences, 7*, 1-12. https://doi.org/10.1186/s13673-017-0116-3.

[28] Manek, A. S., Shenoy, P. D., & Mohan, M. C. (2017). Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier. *World Wide Web, 20*, 135-154. https://doi.org/10.1007/s11280-015-0381-x.

[29] Huq, M. R., Ahmad, A., & Rahman, A. (2017). Sentiment analysis on Twitter data using KNN and SVM. International *Journal of Advanced Computer Science and Applications, 8*(6). https://doi.org/10.14569/IJACSA.2017.080603.

[30] Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*, 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070.

[31] Batistič, S., & van der Laken, P. (2019). History, evolution and future of big data and analytics: a bibliometric analysis of its relationship to performance in organizations. *British Journal of Management, 30*(2), 229-251. https://doi.org/10.1111/1467-8551.12340.

[32] Sharifi, A., Khavarian-Garmsir, A. R., Allam, Z., & Asadzadeh, A. (2023). Progress and prospects in planning: A bibliometric review of literature in Urban Studies and Regional and Urban Planning, 1956–2022. *Progress in Planning, 173*, 100740. https://doi.org/10.1016/j.progress.2023.100740.

[33] Fan, L., Li, L., Ma, Z., Lee, S., Yu, H., & Hemphill, L. (2023). A bibliometric review of large language models research from 2017 to 2023. arXiv preprint arXiv:2304.02020. https://doi.org/10.48550/arXiv.2304.02020.

[34] Farrukh, M., Raza, A., Mansoor, A., Khan, M. S., & Lee, J. W. C. (2023). Trends and patterns in pro-environmental behaviour research: a bibliometric review and research agenda. *Benchmarking: An International Journal, 30*(3), 681-696. https://doi.org/10.1108/BIJ-10-2020-0521.

[35] Wright, R. W., Brand, R. A., Dunn, W., & Spindler, K. P. (2007). How to write a systematic review. *Clinical Orthopaedics and Related Research, 455*, 23-29. https://doi.org/10.1177/0363546513497567.

[36] Yıldırım, E., Mert, K., & Cebeci, H. İ. (2021). Comprehensive Review of the Marketing Ethics Literature: A Bibliometric Approach. *Is Ahlakı Dergisi, 14*(2), 242-269. https://doi.org/10.12711/tjbe.2021.14.2.2059.

[37] Ezugwu, A. E., Shukla, A. K., Agbaje, M. B., Oyelade, O. N., José-García, A., & Agushaka, J. O. (2021). Automatic clustering algorithms: a systematic review and bibliometric analysis of relevant literature. *Neural Computing and Applications, 33*, 6247-6306. https://doi.org/10.1007/s00521-020-05395-4.

[38] Goyal, K., & Kumar, S. (2021). Financial literacy: A systematic review and bibliometric analysis. *International Journal of Consumer Studies, 45*(1), 80-105. https://doi.org/10.1111/ijcs.12605.

[39] Lobonț, O. R., Țăran, A. M., & Costea, F. (2020). E-Government Research Still Matter? A Bibliometric Analysis. *Annals of the University Dunarea de Jos of Galati: Fascicle: I, Economics & Applied Informatics, 26*(2). https://doi.org/10.35219/eai15840409106.

[40] Liu, X., Shi, T., Zhou, G., Liu, M., Yin, Z., Yin, L., & Zheng, W. (2023). Emotion classification for short texts: an improved multi-label method. *Humanities and Social Sciences Communications, 10*(1), 1-9. https://doi.org/10.1057/s41599-023-01816-6.

[41] Zhang, W., Li, X., Deng, Y., Bing, L., & Lam, W. (2022). A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. *IEEE Transactions on Knowledge and Data Engineering, 35*(1), 11019-11038. https://doi.org/10.1109/TKDE.2022.3230975.

[42] Poria, S., Hazarika, D., Majumder, N., & Mihalcea, R. (2020). Beneath the tip of the iceberg: Current challenges and new directions in sentiment analysis research. *IEEE Transactions on Affective Computing, 14*(1), 108-132. https://doi.org/10.1109/TAFFC.2020.3038167.

[43] Hassani, H., & Silva, E. S. (2023). The role of ChatGPT in data science: how ai-assisted conversational interfaces are revolutionizing the field. *Big Data and Cognitive Computing, 7*(2), 62. https://doi.org/10.3390/bdcc7020062.

[44] Hartmann, J., Heitmann, M., Siebert, C., & Schamp, C. (2023). More than a feeling: Accuracy and application of sentiment analysis. *International Journal of Research in Marketing, 40*(1), 75-87. https://doi.org/10.1016/j.ijresmar.2022.05.005.

[45] Hu, H., Xu, J., Liu, M., & Lim, M. K. (2023). Vaccine supply chain management: An intelligent system utilizing blockchain, IoT and machine learning. *Journal of Business Research, 156*, 113480. https://doi.org/10.1016/j.jbusres.2022.113480.

[46] Bouschery, S. G., Blazevic, V., & Piller, F. T. (2023). Augmenting human innovation teams with artificial intelligence: Exploring transformer-based language models. *Journal of Product Innovation Management, 40*(2), 139-153. https://doi.org/10.1111/jpim.12656.

[47] Zhu, T., Li, L., Yang, J., Zhao, S., Liu, H., & Qian, J. (2022). Multimodal sentiment analysis with image-text interaction network. *IEEE Transactions on Multimedia, 25*, 3375-3385. https://doi.org/10.1109/TMM.2022.3160060.

[48] Jin, W., Zhao, B., Zhang, L., Liu, C., & Yu, H. (2023). Back to common sense: Oxford dictionary descriptive knowledge augmentation for aspect-based sentiment analysis. *Information Processing & Management, 60*(3), 103260. https://doi.org/10.1016/j.ipm.2022.103260.

[49] Bello, A., Ng, S. C., & Leung, M. F. (2023). A BERT framework to sentiment analysis of tweets. *Sensors, 23*(1), 506. https://doi.org/10.3390/s23010506.

[50] He, K., Mao, R., Gong, T., Li, C., & Cambria, E. (2023). Meta-based self-training and re-weighting for aspect-based sentiment analysis. *IEEE Transactions on Affective Computing, 14*(3), 1731-1742. https://doi.org/10.1109/TAFFC.2022.3202831.

[51] Liu, M., Luo, X., & Lu, W. Z. (2023). Public perceptions of environmental, social, and governance (ESG) based on social media data: Evidence from China. *Journal of Cleaner Production, 387*, 135840. https://doi.org/10.1016/j.jclepro.2022.135840.

[52] Lian, Z., Liu, B., & Tao, J. (2023). Smin: Semi-supervised multi-modal interaction network for conversational emotion recognition. *IEEE Transactions on Affective Computing, 14*(3), 2415-2429. https://doi.org/10.1109/TAFFC.2022.3141237.

[53] Thakur, N. (2023). Sentiment analysis and text analysis of the public discourse on Twitter about COVID-19 and MPox. *Big Data and Cognitive Computing, 7*(2), 116. https://doi.org/10.3390/bdcc7020116.

[54] Alqarni, A., & Rahman, A. (2023). Arabic Tweets-based Sentiment Analysis to investigate the impact of COVID-19 in KSA: A deep learning approach. *Big Data and Cognitive Computing, 7*(1), 16. https://doi.org/10.3390/bdcc7010016.

[55] Wang, W., Wang, R., Wang, L., Wang, Z., & Ye, A. (2021). Towards a robust deep neural network against adversarial texts: A survey. *IEEE Transactions on Knowledge and Data Engineering, 35*(3), 3159-3179. https://doi.org/10.1109/TKDE.2021.3117608.

[56] Li, R. Y. M., Wang, Q., Zeng, L., & Chen, H. (2023). A study on public perceptions of carbon neutrality in China: Has the idea of ESG been encompassed?. *Frontiers in Environmental Science, 10*, 949959. https://doi.org/10.3389/fenvs.2022.949959.