



HAL
open science

Bibliometric Analysis on the Convergence of Artificial Intelligence and Blockchain

Maryam Hajizadeh, Morteza Alaeddini, Paul Ready

► **To cite this version:**

Maryam Hajizadeh, Morteza Alaeddini, Paul Ready. Bibliometric Analysis on the Convergence of Artificial Intelligence and Blockchain. Blockchain and Applications, 4th International Congress, 595, Springer International Publishing, pp.334-344, 2023, Lecture Notes in Networks and Systems, 10.1007/978-3-031-21229-1_31 . hal-04206522

HAL Id: hal-04206522

<https://hal.science/hal-04206522v1>

Submitted on 13 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bibliometric Analysis on the Convergence of Artificial Intelligence and Blockchain

Maryam Hajizadeh¹[0000-0003-0181-851X], Morteza Alaeddini²[0000-0003-2831-3845], and Paul Reaidy²[0000-0001-8887-4228]

¹ Université Grenoble Alpes, Grenoble IAE, 38000 Grenoble, France
Maryam.Hajizadeh@etu.univ-grenoble-alpes.fr

² Université Grenoble Alpes, Grenoble INP, CERAG, 38000 Grenoble, France
{Morteza.Alaeddini, Paul.Reaidy}@univ-grenoble-alpes.fr

Abstract. Arising from distributed artificial intelligence (AI) and blockchain (BC), the recently emerging notion of decentralized AI refers to transferring information and learning to various peer-to-peer connected machines which learn based on their locally available data and make decisions individually. Using a decentralized consensus mechanism without the need for trustworthy third parties or intermediaries, decentralized AI provides its users with processes, analytics, and decisions based on trusted, digitally signed, and securely shared data that have been transacted and stored on the BC in a decentralized manner. In order to identify core research focusing on both AI and BC and to find avenues for future research, this study employs thematic analysis of keywords of 1,538 academic publications retrieved from WoS and Scopus databases, as well as bibliometric analysis of authors, affiliations, and sources to examine productivity, citation metrics, and bibliographic coupling. By highlighting the areas of digital transformation, environment/society, decentralized AI, DeFi, and cyber security as the focal points of the BC-AI integration, this paper intends to give researchers a comprehensive view of this convergence and can be used in the industry.

Keywords: Machine learning, Distributed ledger, Smart contract, Internet of Things (IoT), Systematic review, Scientometric analysis.

1 Introduction

Due to unending breakthroughs in cutting-edge information technology, perhaps the two growing technologies of artificial intelligence (AI) and blockchain (BC) have played a significant role in driving tremendous changes in most business models in recent years [1]. BC is an immutable, public digital ledger distributed among networked peers [2]. In order to reach consensus, every transaction in the BC is cryptographically signed and validated by all nodes [3], meaning that the nodes keep a copy of the complete ledger consisting of chained blocks of all transactions [4]. Another prominent field of great interest is AI, which allows a computing machine to learn, infer, and adapt depending on its data [4]. The ability of computer-aided machines to perform human behaviors requires skills such as understanding, generalizing, producing solutions,

learning, and inferring from past experiences, which are defined as AI [5,6]

On the one hand, the development of AI has been facilitated by the vast production and creation of data through sensing systems, Internet of things (IoT) devices, social media, and web programs [7]. On the other hand, the ability to configure the BC to manage transactions among people involved in decision-making or data generation and access is a characteristic of BC smart contracts [8]. Like autonomous systems, technologies built based on smart contracts seem to be able to learn and adapt to changes over time and make reliable decisions confirmed and certified by all BC nodes [4]. AI approaches using BC have assisted decentralized learning in analyzing and deciding on vast volumes of securely shared data stored and signed on BC [9].

The research on the combinations of AI and BC is still in the early stages [4]. The idea of recently developing decentralized AI is fundamentally a combination of the two technologies [10]. Because AI works with massive amounts of information, BC now seems to be a secure way to store this data. Focusing on this new approach and aiming to identify potentials and areas of interest for the convergence of two types of digital technologies, BC and AI, our motivation for conducting this research is to depict the status of articles on these two technologies through scientometric visualization.

This paper presents a comprehensive review of the research on BC and AI. Section 2 describes the method we have used to answer the following research questions [11]: (i) Thematically, research in this field can be classified into how many clusters? (ii) Who are the most prominent authors and sources in this field? (iii) Which countries have done the most research in this area? (iv) What are the promising avenues for future research in this field? In Section 3, we answer the questions by co-occurrence mapping of keywords, as well as by providing bibliometric analysis of other attributes of articles. By mapping periodic affiliations and keywords and discussing the research results, in Section 4, we determine several future directions. Section 5 concludes the paper.

2 Method and Data

This work uses a scientometric analysis method to find answers to research questions within the bibliometric attributes of journal articles, conference papers, and book chapters found in the Web of Science (WoS) and Scopus databases using search parameters linked to BC-AI themes. The query string includes the principal terms ‘artificial intelligence’ and ‘blockchain,’ along with their related keywords – computational intelligence, machine learning (ML), neural network, Bayesian network, deep learning, reinforcement learning, supervised learning, Q-learning, transfer learning, natural language processing, thinking computer system, expert system, evolutionary computation, intelligent agent, swarm intelligence, fuzzy logic, hidden Markov model, computer vision, simulated annealing, support vector machine, random forest, decision tree, learning algorithm, genetic algorithm, inductive logic program, intelligent tutoring, autonomous robot, distributed ledger, and smart contract. Also, we only focus on articles in English.

The results include 564 studies from WoS and 1,468 records from Scopus. Finally, we obtained 1,538 studies by eliminating duplicates, merging results, and meeting the criteria (i.e., language and document types). Documents used in this study are from

2008 onwards. The final data set is analyzed using the VOSviewer software tool and the Bibliometrix library in R. Table 1. summarizes the data retrieval resulting from our search strategy.

Table 1. Characteristics of the selected publications.

Data	Documents	Authors
• Retrieval date: 1/26/2022	• Article: 607	• Authors: 4,278
• Databases: WoS; Scopus	• Book chapter: 48	• Author Appearances: 5,919
• Timespan: 2008–2022	• Conference paper: 883	• Authors of single-authored documents: 122
• Sources: 727	• Keywords Plus ^a : 7,352	• Authors of multi-authored documents: 4,156
• Documents: 1,538	• Author's Keywords: 3,345	• Single-authored documents: 132
• Average years from publication: 1.95		• Documents per author: 0.359
• Average citations per document: 6.13		• Authors per Document: 2.78
• Average citations per year per document: 1.55		• Co-authors per document: 3.85
• References: 47,118		• Collaboration Index: 2.96

^a Terms appeared in the titles of an article's references but not appeared in the title of the article itself.

3 Results

3.1 Time Evolution of the Studies

Fig. 1 presents the number of papers in the field of BC-AI published and the mean of total citations between 2008 and early 2022. The number of total publications (TP) over time can be used to track the evolution of research interests.

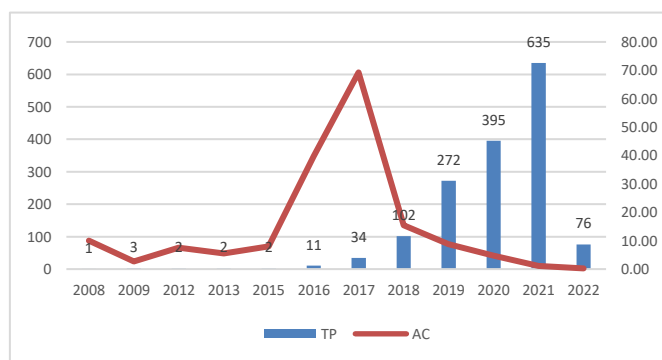


Fig. 1. Annual scientific production in the field of BC-AI.

The annual publication distribution shows an increasing trend. It is observed that BC and AI have become a field of research interest among academics and practitioners in the recent past – only 635 papers were published in 2021. Research in these fields was

limited from 2008 through 2015 as publications are at a stable rate of 1, 2 articles per year. The number of papers rises gradually to 11 and 34 in 2016 and 2017, then increases sharply from 102 in 2018 to 635 in 2021. Based on the trend of recent years and the actual number of publications in the first month of 2022, approximately 1,066 publications are projected for 2022. Also, the average citations (AC) rate per year in Fig 1 shows that the articles published in 2016 and 2017 have been cited more than others.

3.2 Thematics of the Studies

In order to identify important research topics in BC and AI, we first analyze authors' keywords that represent the vital content of the studies. The results confirm that smart contract, security, and IoT are at the top of the BC-AI keywords with 459, 355 and 321 occurrences, respectively. In order to remove general and merge similar terms, we build a thesaurus and use it in keyword mapping. To calculate the relatedness of the keywords for building a co-occurrence map, we employ the fractional-counting method, in which the weight of a co-occurrence link is fractionalized based on the number of other keywords in the publication [12]. We set at least 30 occurrences of keywords, meaning we cluster keywords that occur in more than 2% of documents. With these settings, 42 keywords meet the threshold.

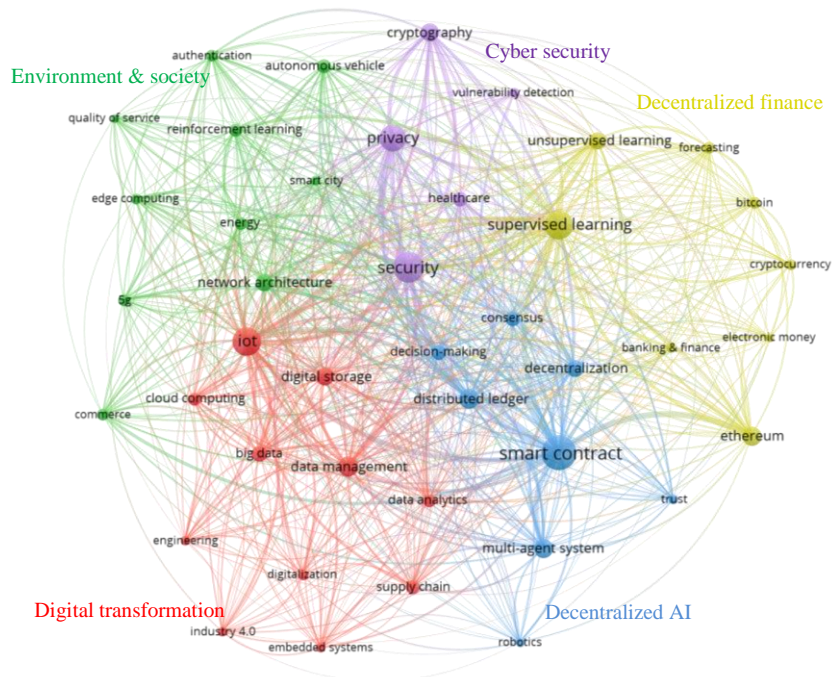


Fig. 2. Clusters resulting from the keywords co-occurrence analysis.

The network depicted in Fig. 2 is composed of five clusters of keywords. We name clusters according to the meaning and nature of most of the keywords belonging to

them. The red cluster, digital transformation, includes keywords of digital technologies that are primarily used to meet changing business and market requirements. This cluster reflects the authors' focus on the combined use of AI and BC in data, cloud, and IoT as well as their application in industry and the supply chain. The green cluster, environment & society, highlights the application of AI and BC to topics such as energy, smart cities and vehicles, commerce, and other fields of interest to society and the environment. The exciting thing about this cluster is that researchers use reinforcement learning algorithms to develop intelligent agents who perceive and interpret their environment, take actions and learn through trial and error. The other three clusters in Fig. 2 refer to the use of blockchain to provide reliable data for robots and intelligent agents to improve the quality of their decisions (blue), predictions in the field of decentralized financial (DeFi) using both supervised and unsupervised learning algorithms (yellow), and concepts regarding the protection of blockchain network against cyber attacks (purple). The nodes or keywords that did not have a network with other keywords have the potential to become new research topics in the future [13] (see Section 4).

3.3 The Most Prolific Authors

Fig. 3 shows in three ways the top ten authors who have the highest publication productivity in BC-AI. First, the number of publications that appear in the author's name (as the lead author or co-author) has been counted. The total citations of each author are presented in the second measure, and the third one indicates the authors' h -index, showing that authors have authored h papers, each of which has been referenced at least h times. From Fig. 3(i), Wang Y and Wang X are ranked first with 22 articles, followed by Li Y and Li J with 20 papers and Wang Z and Zhang Y with 18 articles. Although some authors are present in both Figs. 3(i) and 3(iii), in Fig. 3(ii), we come across several names that do not exist in those two figures. For instance, Wang X, Wang Y, Li J, Tanwar S, and Wang Z, who have high TP ratings, also have the highest h -index. In Fig. 3(ii), Atzei N, Bartoletti M, and Cimoli T are co-authors of an article with 649 citations which put them in the first citation rank.

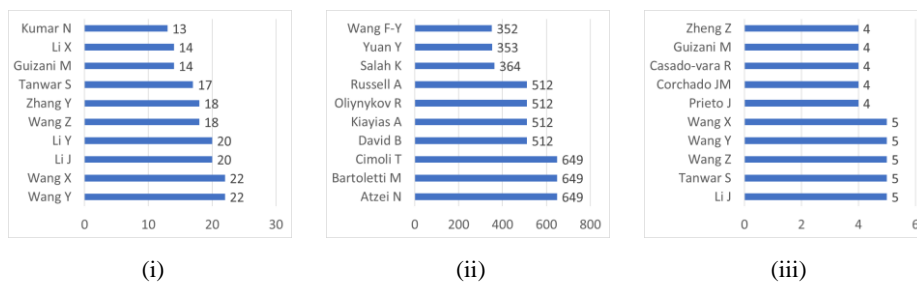


Fig. 3. Authors' productivity: (i) TP-total publications, (ii) TC-total citations, (iii) h -index.

To visualize the collaboration of authors with each other, we make a co-authorship map when the minimum numbers of publications and citations are 5 and 0, respectively (Fig. 4). Intensely collaborative authors from the red cluster include Li J, Li X, and Zhang J, who study BC, ML, Ethereum, and multi-agent systems. The green cluster

involves the authors studying BC, ML, IoT, and federated learning. Wang Y and Wang X are the most productive authors in this cluster. Tanwar S, Kumar N, and Gupta R emerge as solid contributors in AI and distributed ledgers in the blue cluster. In the yellow clusters, Chen Y and Wang S are the most collaborative authors whose main fields of interest are BC and privacy issues. Topics proposed in the violet cluster are BC and security, in which Guizani M and Chamola V have the maximum productivity.

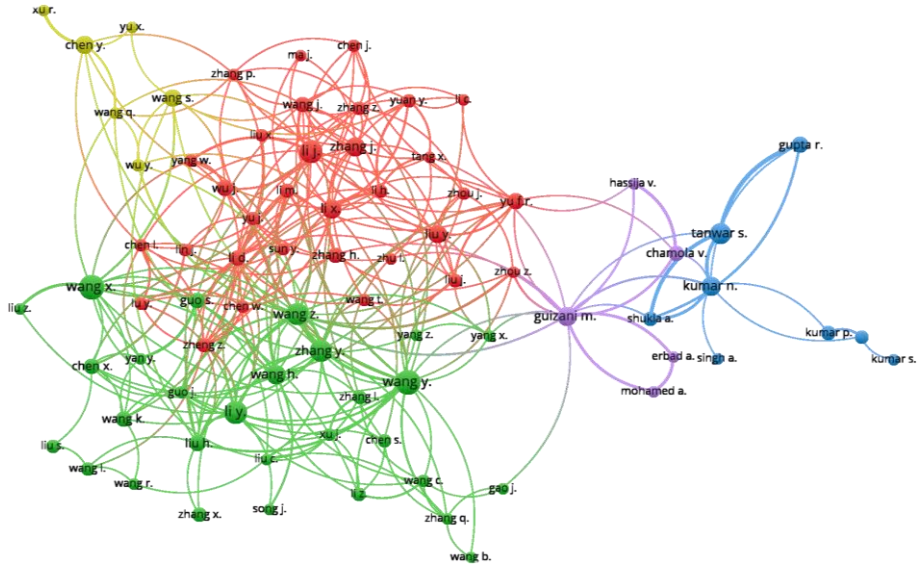


Fig. 4. The co-authorship network.

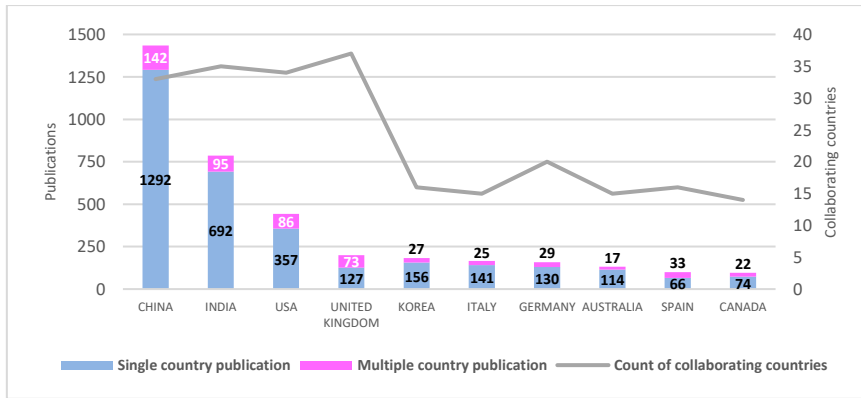


Fig. 5. Corresponding authors' countries.

3.4 Geographic Collaborations Perspective

The authors of the documents reviewed in this study are from 81 different countries. Fig. 5 Shows 10 countries with the highest number of the identified papers according

to the geographical locations of the authors. On top of all countries, China is the highest producer of research in this field with 1,434 studies, of which 1,292 documents have been published by Chinese researchers, and 142 publications have been prepared in collaboration with researchers from 33 other countries. In this regard, China is followed by India with 787 publications (692 publications alone and 95 publications in collaboration with 35 countries) and the United States with 443 documents (357 single publications and 86 joint publications with 34 countries). The results of our study show that the term blockchain is the main keyword in about half of China’s research, followed by the term smart contract. The situation is similar in India and the United States, although the weight of focusing on these keywords may be somewhat different.

3.5 The Most Influential Sources

In clustering by coupling sources measuring and cluster labeling by authors’ keywords with 500 units and 5% minimum frequency, we have two clusters: the first consists of smart contracts and ML, and the second covers IoT, security, and privacy. The top three publication venues in terms of total papers and total citations are shown in Table 2. By publishing 114 articles, Lecture Notes in Computer Science is the most popular series of distinguished conference proceedings in the field of BC-AI. To measure the impact of each source on this research community, we also evaluate those shortlisted publication venues based on the total number of citations; the results show that Lecture Notes in Computer Science has outperformed others in terms of the total number of citations.

Table 2. Source impact: TP–total publications, TC–total citations.

Source	TP	Source	TC
Lecture Notes in Computer Science	114	Lecture Notes in Computer Science	1,518
IEEE Access	50	IEEE Access	936
Advances in Intelligent Systems and Computing	46	IEEE Internet of Things Journal	540

4 Discussion and Future Research Direction

Fig. 6 presents the thematic evolution of BC-AI research based on the authors’ use of keywords over time. From 2008 to 2015, most topics revolve around cryptography and the use of supervised learning. Gradually, other themes evolve in the 2016–2019 period due to spreading technology – AI methods (e.g., fuzzy logic and unsupervised learning), BC topics (e.g., smart contracts, distributed ledger, and consensus), IoT, privacy, and e-commerce. Interestingly, research focused on cryptography and supervised learning in the first era are mainly used in IoT field in the second period. The period of 2020–2022 witnesses research in IoT, security, and BC-related topics. Moreover, smart city and reinforcement learning are the new topics that have been addressed in this period. It seems that these topics, as well as cryptography and supervised learning methods that have returned to the field of interest of scholars, remain for future development.

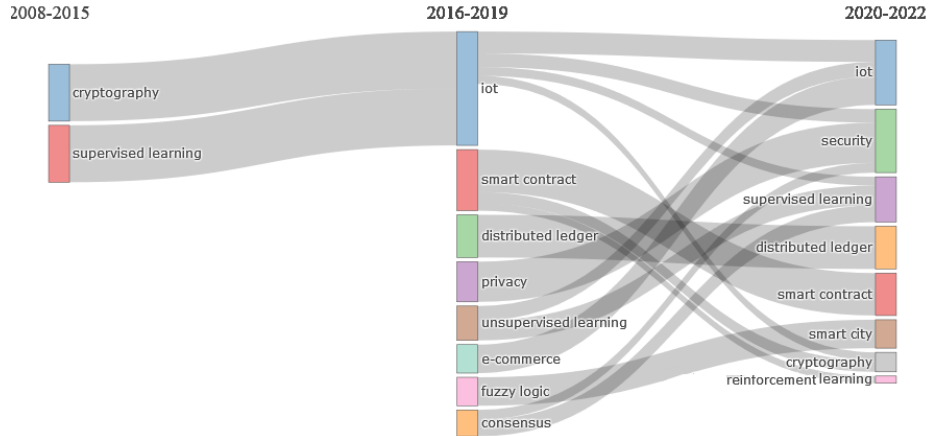


Fig. 6. Multi-fields plot between affiliation and keywords.

The co-occurrence map and the most frequent word analysis suggest that extant literature has made significant strides in understanding the BC and AI concepts and their application in the fields of IoT, security, privacy, supply chain, healthcare, and finance. Although research on these sectors seems mature, there are several gaps and dimensions that should be explored in the future. Managing and controlling localized data by fog nodes equipped with AI, ML and BC is an example.

Although BC and AI have advanced considerably in recent years, there remains a significant gap in terms of combining these two concepts. While empirical investigations of this combination have already highlighted a few possible benefits of this integration, future research could help further converge these technologies. Both technologies have distinct advantages that, if combined, will promise more reinforced benefits that would cause the applications of both fields to grow.

Research into different literature under this bibliographic analysis reveals several common avenues for future research. This includes research on their vulnerability, how they help develop autonomous renewable energy, their application in the circular economy, investigating the result of the combination of federated learning protocols with BC, or using game-theoretic mechanisms to improve cryptographic technologies for secure computation. Work on privacy-enhancing technologies (PETs) is one of the fields that has recently prompted scholars to study this integration. Recent changes to PETs using blockchain and smart contracts that focus on differential privacy (e.g., PETchain by Javed et al. [14]), federated learning (e.g., FLchain by Nguyen et al. [15]), and privacy-preserving (e.g., Deebak and Fadi [16]) are examples confirming garnering interest in this convergence. Establishing a mechanism for distributed trust among nodes in a BC network who use AI to decide whether to endorse other nodes seems to be another direction for future research in this area. Moreover, BC technologies have shown high adherence in operations and supply chain management, while studies contemplating the interplay of BC, AI, and supply chain are in the early stages.

5 Conclusion

In this paper, an evaluation of the global research trends in BC-AI publications from 2008 to 2022 is given. The convergence of BC-AI and their related topics has been a field of extensive research during the last five years, and its publication output is characterized by exponential growth. The latest trends indicate that scientific output in this field is expected to grow. The next few years will enlighten if this rising trend in publication output will remain. The study includes 1,538 publications on BC-AI, covering 4,278 authors, 727 publication venues, and 81 countries.

Five leading research areas can be distinguished in the domain of BC-AI: (i) digital transformation, (ii) environment and society, (iii) decentralized AI, (iv) DeFi, and (v) cyber security. The research area of DeFi seems to be more theoretical, whereas the research areas of digital transformation and cyber security have a more practical-oriented emphasis. The research in the subdomains of digital transformation, environment, and society stands in a dominant position in BC-AI research nowadays.

Some favorable aspects derived from the bibliometric analysis are as follows: (i) Since multi-authored publications make up more than 90% of all publications, it can be concluded that there is much collaborative research in the field of BC-AI; (ii) The fact that many different sources with a wide range of subject categories are devoted to BC-AI publications indicates the wide variety of research themes and the multidisciplinary character of the research in this domain.

China, India, and the United States are the most influential countries in this field. Based on the highest number of publications, Beijing University has the most heightened cooperation and activities. Moreover, according to the number of citations and papers, Lecture Notes in Computer Science have the highest impact among sources. Furthermore, although Wang Y and Wang X rank first in publishing, Atzei N, with “A survey of attacks on Ethereum smart contracts,” is the most influential among authors. Also, our study discovers the potential future research in BC and AI, including a combination of federated learning protocols with BC, a distributed trust mechanism using BC and AI, and practical research to use BC and AI in supply chain management and in the fields of security and privacy.

Some limitations of this scientometric study are as follows. First of all, the search has been limited to publications listed on WoS and Scopus. Although they are among the most prominent global databases, they do not contain all publications in the field of BC-AI research. Other international databases such as ProQuest could have been used. It should also be noted that the integration of Scopus and WoS databases and the analysis of data within the final unique repository with VOSviewer and R are associated with problems and differences in their reports. The second limitation is the predominance of quantitative methods in bibliometric analysis, making the content or quality of publications uninterpretable [17]. This could lead to some publications being included in the analyses despite addressing a different topic than BC-AI convergence.

Finally, we hope that this bibliometric analysis of the confluence of AI and BC can give crucial information to the researchers and help them better understand the evolution of technical ecosystems and the growing literature in this cross-disciplinary field. Keeping up with the rapidly developing body of BC-AI research allows academics and

practitioners to take advantage of the opportunities created by the confluence of the two domains to promote the use of AI for BC and BC for AI.

References

1. Singh SK, Rathore S, Park JH (2020) Blockiotintelligence: A blockchain-enabled intelligent IoT architecture with artificial intelligence. *Future Generation Computer Systems* 110:721-743
2. Nakamoto S (2008) Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review*:21260
3. Alaeddini M, Dugdale J, Reaidy P, Madiès P, Gürcan Ö (2021) An Agent-Oriented, Blockchain-Based Design of the Interbank Money Market Trading System. In: *Agents and Multi-Agent Systems: Technologies and Applications 2021*. Springer, pp 3-16
4. Salah K, Rehman MHU, Nizamuddin N, Al-Fuqaha A (2019) Blockchain for AI: Review and open research challenges. *IEEE Access* 7:10127-10149
5. Bello O, Holzmann J, Yaqoob T, Teodoriu C (2015) Application of artificial intelligence methods in drilling system design and operations: a review of the state of the art. *Journal of Artificial Intelligence and Soft Computing Research* 5
6. Kalogirou SA (2003) Artificial intelligence for the modeling and control of combustion processes: a review. *PrECS* 29(6):515-566
7. Schmidhuber J (2015) Deep learning in neural networks: An overview. *NN* 61:85-117
8. Wood G (2014) Ethereum: A secure decentralised generalised transaction ledger. *Ethereum project yellow paper* 151(2014):1-32
9. McMahan B, Moore E, Ramage D, Hampson S, y Arcas BA Communication-efficient learning of deep networks from decentralized data. In: *Artificial intelligence and statistics, 2017*. PMLR, pp 1273-1282
10. Hussain AA, Al-Turjman F (2021) Artificial intelligence and blockchain: A review. *Transactions on Emerging Telecommunications Technologies* 32(9):e4268
11. Mejia C, Wu M, Zhang Y, Kajikawa Y (2021) Exploring Topics in Bibliometric Research Through Citation Networks and Semantic Analysis. *Frontiers in research metrics and analytics* 6
12. van Eck NJ, Waltman L (2014) Visualizing bibliometric networks. In: *Measuring scholarly impact*. Springer, pp 285-320
13. Effendi DN, Anggraini W, Jatmiko A, Rahmayanti H, Ichsan IZ, Rahman MM Bibliometric analysis of scientific literacy using VOS viewer: Analysis of science education. In: *Journal of Physics: Conference Series, 2021. vol 1*. IOP Publishing, p 012096
14. Javed IT, Alharbi F, Margaria T, Crespi N, Qureshi KN (2021) PETchain: a blockchain-based privacy enhancing technology. *IEEE Access* 9:41129-41143
15. Nguyen DC, Pathirana PN, Ding M, Seneviratne A (2020) Integration of blockchain and cloud of things: Architecture, applications and challenges. *IEEE Communications surveys & tutorials* 22(4):2521-2549
16. Deebak BD, Fadi A-T (2021) Privacy-preserving in smart contracts using blockchain and artificial intelligence for cyber risk measurements. *Journal of Information Security and Applications* 58:102749
17. Pajo AT, Espiritu AI, Jamora RDG (2020) Scientific impact of movement disorders research from Southeast Asia: a bibliometric analysis. *Parkinsonism Relat Disord* 81:205-212