Scientometric Analysis in the Field of Big Data and Artificial Intelligence in Industry

Fairuz Iqbal Maulana
Computer Science Department,
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
fairuz.maulana@binus.edu

M. Aldiki Febriantono
Computer Science Department,
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
m.aldiki@binus.ac.id

Miftahul Hamim
Entrepreneurship Business Creation,
BINUS Business School
Bina Nusantara University
Jakarta, Indonesia 11480
miftahul.hamim@binus.edu

Bayu Ramadhani Fajri
Department of Animation,
Faculty of Engineering,
Universitas Negeri Padang,
Padang, Indonesia
bayurf@unp.ac.id

Rahman Arifuddin
Department of Electrical Engineering
University of Merdeka Malang
Malang, Indonesia
rahman.arifuddin@unmer.ac.id

2022 1st International Conference on Information System & Information Technology (ICISIT) | 978-1-6654-0200-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICISIT54091.2022.9872659

***Abstract*— Big Data and Artificial Intelligence (BD&AI) in Industry have grown so prevalent, and the potential they provide is so revolutionary that they are seen as critical for competitive growth. Because the number of organizations BD&AI on Industry technology is increasing exponentially, so is the need for BD&AI on Industry practitioners. Until we conducted this research, only 1399 academic documents on BD&AI in Industry found from 2002 to 2020 were obtained by searching the Scopus database. BD&AI in the industrial sector is examined in-depth in this paper. This study uses bibliometric analysis and indexed digital methods to map scientific publications worldwide. This study uses the Scopus database to collect information and online analysis via the Scopus website and VOSViewer to demonstrate bibliometric network mapping. We use the article selection process, starting with the keywords to be searched for, the year limitation, then the database is exported into RIS and CSV format files. From the database, we also perform network mapping using VOSViewer. Researchers in China have the most articles published and indexed by Scopus among the most prolific authors (373), followed by the United States (239) and India with 125 academic publications. Data analysis reveals an upward trend in the number of worldwide publications in BD&AI in Industry, as measured by the Scopus index.**

Keywords—big data, artificial intelligence, industry, bibliometric, VOSViewer

I. INTRODUCTION

The term "Big Data" refers to the collection, storage, and analysis of massive amounts of data using a variety of technologies, individuals, and procedures. Every second, massive amounts of data are created from sources all around us. Businesses may improve their operational flexibility [1] and adaptability by translating this data into valuable information using Big Data (BD) technology such as AI (Artificial Intelligence) [2]. In general, artificial intelligence (AI) refers to interdisciplinary attempts to make computers and systems intelligent enough to perform efficiently and predictably. Machine learning (ML), Computer Vision (CV), robotics, and NLP (Natural Language Processing) have laid the groundwork for artificial intelligence applications during the past few decades. The worldwide economic effect of AI is estimated to exceed $15.7 trillion by 2030, and AI capability development has been designated as a national priority in the United States. Artificial Intelligence (AI) can be used to gain new insights, improve decision-making flexibility and generate new kinds of value in computational large data analysis. With the advent of Big Data and Artificial Intelligence (AI), businesses will need to become more agile and competitive in order to remain relevant. [3].

A number of developed economies have implemented regulations to encourage the use of Big Data and Artificial Intelligence technologies, notably the European Union (EU) and the United States of America (US). Data value chain and the data-driven economy is a top focus for the European Commission's partnership with European companies, researchers and institutions. Horizon Europe is Europe's largest research and innovation effort that provides over 95.5 billion Euros in funding towards the development of Big Data and Artificial Intelligence technologies, for example [4]. It was developed by the United States in 2012 to improve Big Data and Artificial Intelligence technology as well as train the next generation of data scientists and analysts. For their part, 15 federal agencies developed a strategy under the leadership of the Big Data Senior Steering Group that included initiatives including building a collaborative Big Data innovation ecosystem, modernizing Big Data's digital infrastructure, and adopting Big Data education and training. The White House gave its blessing to the scheme.

Parallel to these projects, there is a continuous discussion about how Big Data and Artificial Intelligence transforms industrial areas such as life sciences, finance, manufacturing, marketing, logistics, and customer service [5]. Applied to the life sciences industry, Big Data and Artificial Intelligence help to enhance medicine discovery processes, treatment ideas, patient involvement and adherence, as well as administrative tasks (e.g., healthcare) [6]–[8]. There are various possible hurdles identified in this research, including as data security and privacy problems as well as ethical considerations, that might hinder large-scale process automation in the medical sciences business. To transform traditional manufacturing facilities into ideal smart manufacturing facilities, Big Data and Artificial Intelligence are of great importance, according to another study (e.g., Cloud Computing, Cyber-Physical Systems, and Internet of Things) [9]–[11]. A couple of the research's main takeaways are automation of manufacturing processes (from production to distribution) and acceleration of the product development life cycle. In today's manufacturing world, both of these discoveries underline the necessity of Big Data and Artificial Intelligence. Another researcher examines the impact of big data and artificial intelligence on the financial and

management accounting industries, as well as reporting services [12]. BD&AI can be used to detect false accounting statements, predict stock market performance, generate new accounting standards, and manage financial risk. According another research, The marketing and customer service sectors are likewise being transformed by BD&AI [13]. It is the focus of these articles to highlight the advantages of using video, audio, and even text analytics to uncover client demands and anticipate future behavior.

In addition to studies on the impact of BD&AI on various industries, papers have been written on the impact of BD&AI on the labor market. For example, it has been noted that the present workforce does not have the skills required by the job market [14]. Information technology workers are desperately needed in the workforce, according to the authors. Another researcher [15] present a paradigm for identifying a flexible workforce enablers in BD&AI age, with a special focus on IoT. Beilby [16] investigated how Big Data And Artificial Intelligence affects the healthcare profession and suggests solutions for training doctors with IT skills to use BD&AI tools for clinical decision making. Some researcher [17] used degrees of automation to determine the digital, and soft skills of various vocations.

The bibliometric or scientometric analysis is used to process the Scopus database from the search results of articles in scientific publications on research on Big Data and Artificial Intelligence (BD&AI) topics in Industry. Using the Scopus database, we analyzed research published from 2001 to 2020 to see where researchers have focused their efforts. The first step is to limit article search keywords with "Big Data and Artificial Intelligence on Industry". The second step specifies the year of publication as needed. Then, the data is exported into RIS and CSV files for mapping analysis purposes using VOSViewer software. There are 1399 scientific documents on BD&AI in Industry with indexed by Scopus. Scopus visual maps and VOSViewer were used to process and visualize data, which was then analyzed bibliometrically.

II. Research Methods

A systematic and clear bibliometric review technique that focuses on the limitations of knowledge [18]. For the purposes of bibliometric analysis, the search and selection criteria are shown in Figure 1. Scopus, one of the most extensively cited archives, was used to acquire the data. Search terms for this study were "Big Data and Artificial Intelligence on Industry" to avoid picking papers that were irrelevant to the study's purpose. On December 13, 2021, a search was conducted. Total retrieved papers were 1959, but this was reduced to 1399 when publications from 2021 and 2022 were excluded from the analysis. A total of 1399 publications were studied for their bibliometric data. The procedure of selecting an article is shown in Figure 1.

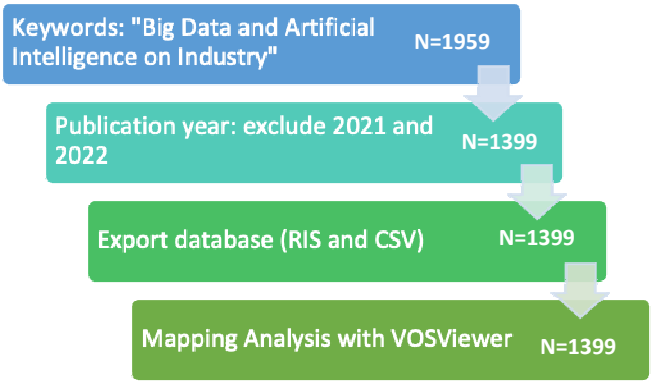

Fig. 1. Article selection process

III. Result and Discussion

A. Most Document of BD&AI on Industry Publication by Country

This is seen in Figure 2: Document comparison for up to ten nations in BD&AI on Industry publications. China is the country with the most BD&AI on Industry publications, with 373. Then, the United States with 239 documents research papers, India has 125 documents research papers, the United Kingdom has 73 documents research papers, South Korea has 70 documents research papers, Italy has 54 documents research papers, Germany has 50 documents research papers, Spain has 47 documents research papers, Australia has 50 documents research papers, and the Russian Federation has 50 document publications.

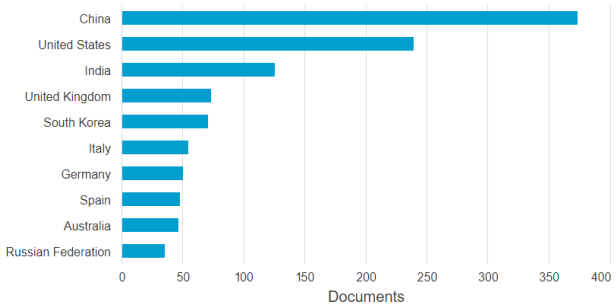

Fig. 2. Compare the number of documents from up to ten different nations

B. Most Frequent Institution Affiliation of BD&AI on Industry Publication

Compare the number of documents for up to ten connections in BD&AI on Industry publications, as shown in Figure 3 are the Amity University have 14 documents research papers, start from 2015-2020. Beihang University is represented by the second chart line, which contains 12 papers spanning the years 2016 through 2020. In 2017, 411 people cited this document. Thirdly, Shanghai Jiao Tong University has 11 documents of connection with the institution.

After Hong Kong Polytechnic University, China's Ministry of Education have 8 document affiliations, and Peter the Great St. Petersburg Polytechnic University have 8 document affiliations too, the next line reads: Seven documents each from the University of Johannesburg,

Zhejiang University, the State Grid Corporation of China, and Sichuan University.

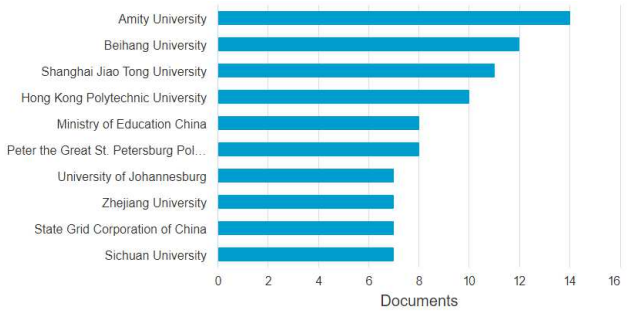

Fig. 3. Up to ten affiliations' document counts can be compared

C. The Majority of Authors in BD&AI on Industry Publications are Single Authors

In terms of industry publications, the three writers with the most Big Data and AI articles are Tao, F., with five documents, then the second line is Huh, J. H., and Perera, L. P., with four papers each author. The next line author is Calvo, P., Chen, Y., Cheng, X., Cheong. H. W., Engkvist, O., Lammer, G., Lee .H, have three authors' documents.

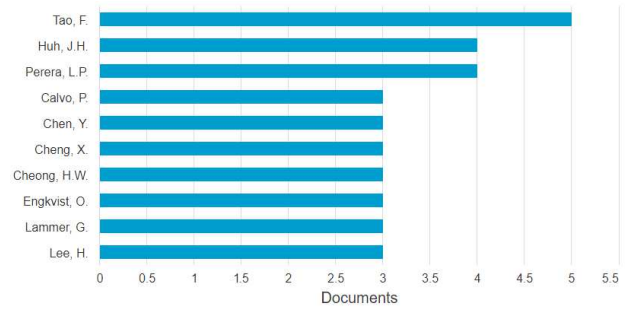

Fig. 4. Up to ten authors' document counts can be compared

Tao, Fei from Beihang University, Beijing, China also collaborated with several other researchers. This demonstrates the existence of interdisciplinary research collaboration.

D. Most Frequency of BD&AI on Industry Publication by Subject Area

There is found 1400 documents result of BD&AI on Industry publication from 1999 until 2020. The most knowledge in the area of Computer Science subject area with 830 academic documents (28.9%). The second topic area is Engineering with 555 academic documents (19.3%), and the third is Mathematics with 217 academic documents (7.6%). Decision Sciences subject area have 199 documents (6.9%), Business, Management and Accounting subject area have 145 documents (5.1%), Social Sciences subject area have 134 documents (4.7%), Medicine subject area have 123 documents (4.3%), Energy subject area have 116 documents (4.0%), Physics and Astronomy subject area have 110 documents (3.8%), Materials Science subject area have 82 documents (2.9%), and other with 12.5% like figure "Documents of BD&AI on Industry Publication by Subject Area" (Fig.5).

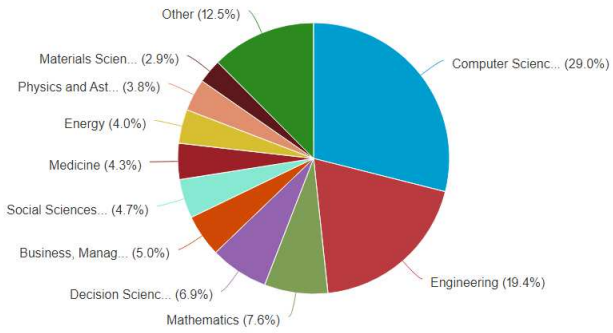

Fig. 5. Documents of BD&AI on Industry Publication by Subject Area

E. The most common type of Big Data and AI (BD&AI) document on Industry Publication.

Scholars have published a wide range of scholarly papers on the subject of BD&AI on Industry, including eleven distinct categories. Figures from the diagram show that BD&AI on Industry academic document publication shows Conference Paper publications (57.1 percent) with 810 academic documents, Article (27.6 percent), Review (6.4 percent), Book Chapter (3.9 percent), Conference Review (1.6 percent) with 22 documents, Book (1.4 percent) with 19 document, Editorial (0.5 percent) with 7 documents, and Note (0.4 percent) with 4 academic documents, respectively, like in the figure chart of BD&AI publication types (Fig.6). There are two main publication types with the highes document by type: Conference Papers and Articles.

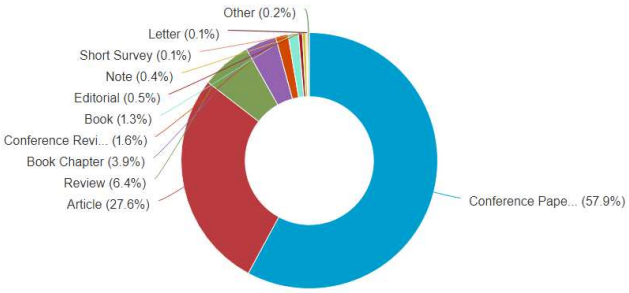

Fig. 6. Chart of BD&AI Publication Types

F. Big Data and Artificial Intelligence on Industry Publications Document Per Year

According to worldwide publication sources, below are the top five documents published each year. The number of articles published each year in the field of Big Data and Artificial Intelligence on Industry is indicated by "Journal Of Physics Conference Series" with 45 documents from 2018 to 2020 in the purple line color, followed by "Lecture Notes In Computer Science Including Subseries Lecture Notes In Artificial Intelligence And Lecture Notes In Bioinformatics" with 43 docs, then Figure 7 shows a graphical representation of the situation.

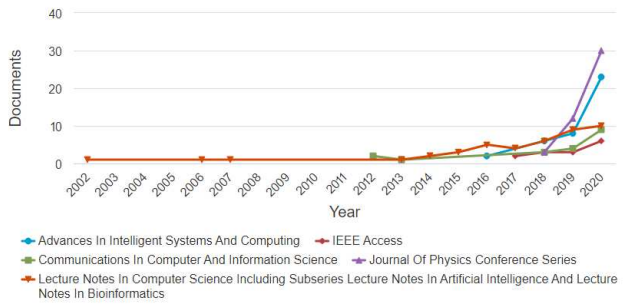

Fig. 7. Big Data and Artificial Intelligence on Industry Publication Documents Per Year Sources Graph

Researchers in BD&AI on Industry publications prefer publishing output to Proceedings, as evidenced by the aforementioned statistics. According to experts, the amount of pages in contrast to publications may be limited. It is projected that research quality will continue to improve in the near future.

G. Documents per year of BD&AI on Industry Publication

A rise in the number of BD&AI on Industry articles produced annually between 1999 until 2020 may be seen. In 1999 until 2012, just few papers were released; seems only 1 until 6 documents. But, in 2013 that number had risen to 9 documents. Then start to increase more and more until 2020 with 576 documents.

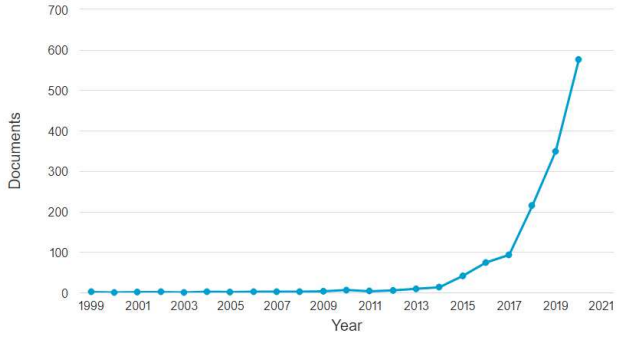

Fig. 8. BD&AI on Industry Literature Documentation Yearly Graph by Source

H. Documents most cited from Big Data and Artificial Intelligence (BD&AI) on Industry Publication

In 2008, Computers and Education published the most referenced international paper on the topic of BD&AI on Industry. The top three most-cited papers from each year are included in the table below:

TABLE I. THE TOP THREE MOST CITED PUBLICATIONS IN BD&AI ON INDUSTRY

No.	BD&AI in Industry Documents Scopus Publication				
	Document Title	***Authors***	***Year***	***Source***	***Cited by***
1.	"Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing"[19]	Tao, F., Zhang, M.	2017	IEEE Access 5,8049520, pp. 20418-20427	414
2.	"Brave new world: service robots in the frontline"[20]	Wirtz, J., Patterson, P.G., Kunz, W.H., (...), Paluch, S., Martins, A.	2018	Journal of Service Management 29(5), pp. 907-931	344
3.	"Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges"[21]	Wang, Y., Chen, Q., Hong, T., Kang, C.	2019	IEEE Transactions on Smart Grid 10(3),8322199, pp. 3125-3148	335

I. Map of Publication Theme from BD&AI in Industry

The keyword network mapping for BD&AI in Industry was carried out with the help of the VOSViewer program. These are the fields that we will be using for this research: title and abstract. The computation is carried out using binary counting. The bare minimum of keyword-related documents is two occurrences per keyword-related document. 718 terms are eligible for consideration out of a total of 29462 terms. In this case, the phrase must appear at least 10 times in the sentence. There will be a score assigned to each of the 718 words based on their relevance. The default option is to select the phrase that is 60% most relevant, and the total number of phrases to be selected is 431, as shown in the picture below:

Selected	Term	Occurrences	Relevance
☑	support vector machine	17	4.23
☑	deep neural network	11	4.22
☑	special focus	13	3.49
☑	random forest	16	3.41
☑	decision tree	19	3.31
☑	upgrading	15	3.01
☑	design methodology approach	14	3.01
☑	originality value	15	2.97
☑	proceeding	23	2.87
☑	fintech	12	2.86
☑	temperature	17	2.73
☑	speech recognition	12	2.56
☑	paper analyze	13	2.54
☑	novel approach	11	2.36
☑	lstm	15	2.35
☑	genetic algorithm	11	2.33
☑	conference	31	2.29
☑	ai technique	11	2.28
☑	reform	19	2.28
☑	gas industry	20	2.19
☑	drug discovery	10	2.18

Fig. 9. Verify selected terms of BD&AI on Industry Publication

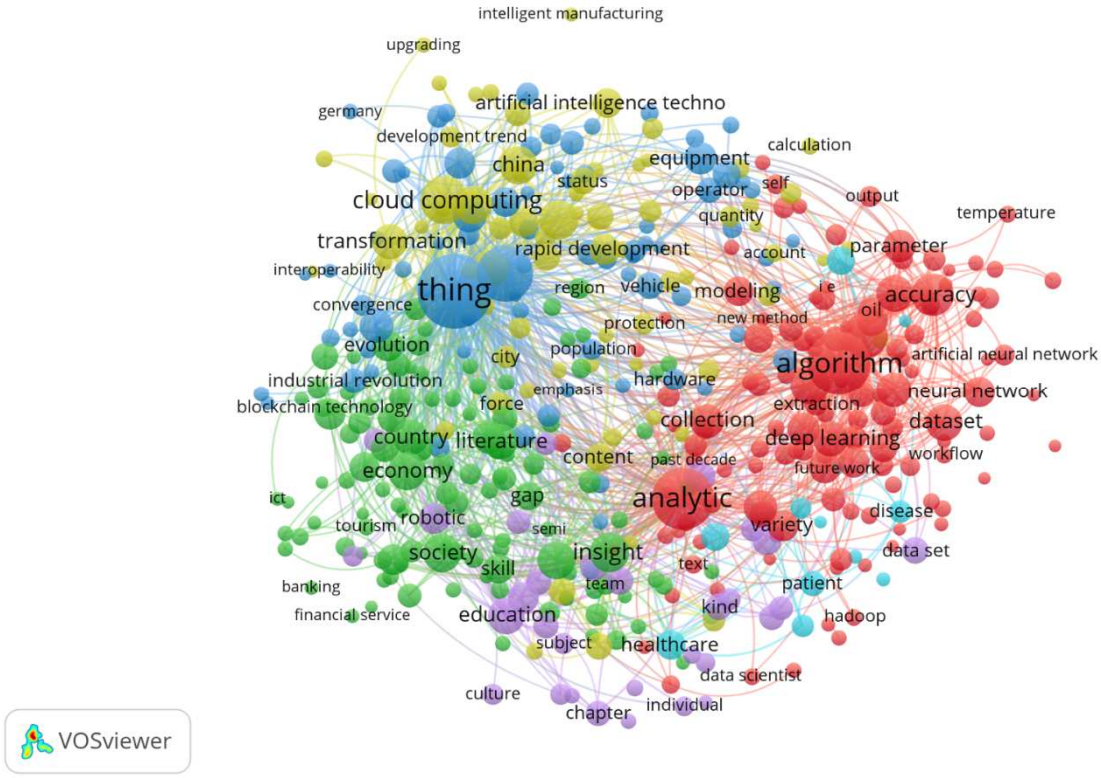

Fig. 10. The coexistence of 431 of the most common terms (with at least five occurrences). The thickness of the lines shows the intensity of the association between keywords, which was assessed by the frequency with which they occurred in published articles together. VOSViewer-generated artwork

Figure 10 depicts two groupings of keyword-based research themes:

1. Algorithm (red). This group includes 127 items featuring the term network like academia., analytic, ai application, ai system, ai technique, artificial intelligence technology, artificial neural network, big data processing, business intelligence, cluster, classification, collection, computation, computer science, deep learning, decision tree, deep neural network, engineering, modelling, and others item.
2. Blockchain (green). This group includes 113 items featuring the term network like literature, academic, big data technique, blockchain, business model, conference, cyber security, data security, employee, edge, fintech, ict, manager, literature review, marketing, and others item.
3. Thing (blue). This group includes 75 items featuring the term network like iot, asset, big data analytics, building, disruptive technology, driver, date, failure, hardware, historical data, maintenance, manufacturer, key role, networking, operator, procedure, realization, and others item.
4. Cloud Computing (yellow). This group includes 72 items featuring the term network like account, ai technology, analyze, calculation, content, course, feasibility, fusion, future development, further development, development trend, direction, feasibility, and others item.
5. Education (purple This group includes 30 items featuring the term network like art, augmented reality, author, beginning, chapter, culture, data set, education, finance, logistic, medicine, search, institution, price, property, robotic, search, subject, training, and others item.
6. Education (purple). This group includes 14 items featuring the term network like current state, decision marker, decision support system, diagnosis, disease, dss, future work, health care, healthcare industry, healthcare sector, hospital, patient, and treatment.

J. Research's Author Collaboration Network

A trend of collaborative research may be seen in figure 11 of a Big Data and Artificial Intelligence on Industry article. The Author Network of Collaboration has 1731 writers, three of whom meet the condition of having written at least four documents. Then choose three authors from the pool. There are three distinct research teams, each of which is related to the others.

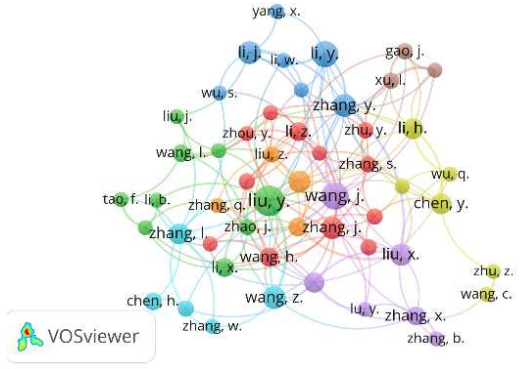

Fig. 11. The network includes 55 co-authors who write in the fields of BD&AI on Industry. Each node features a different author. The number of publications is represented by the size of the nodes, while the level of collaboration is shown by the thickness of the lines between nodes. VOSViewer was used to create the artwork.

IV. Conclusion

The findings of this research point to patterns and trends in the increasing number of Scopus-indexed international papers published in the journal "Big Data and Artificial Intelligence on Industry." China has the most materials on BD&AI in Industry publications, with 373. The United States has 239 document publications, followed by India, 125 document publications. Computer science dominates BD&AI in the Industry sector, with 830 scholarly articles (29.0%). The second, Engineering, has 555 educational documents (19.4%). The third sector is Mathematics, with 217 academic papers (7.6%).

This study provides a classification of BD&AI in the Industry convergence axis, based on the contribution to knowledge made by the companies involved. BD&AI in Industry research may have to wait several years to uncover the key issues. More topics may be examined or investigated, allowing for a complete comprehension of the subject matter at hand. Using a combination of Scopus, Web of Science, and other indexations, the contribution and effect of Big Data and Artificial Intelligence on Industry should be evaluated as soon as feasible in the future.

Acknowledgment

Please accept our heartfelt thanks on behalf of our family and friends, particularly those on the BINUS Malang campus and at other colleges, for your assistance in completing this project on time. We appreciate your assistance.

References

[1] S. K. Shukla, Sushil, and M. K. Sharma, "Managerial Paradox Toward Flexibility: Emergent Views Using Thematic Analysis of Literature," *Glob. J. Flex. Syst. Manag.*, vol. 20, no. 4, pp. 349–370, 2019, doi: 10.1007/s40171-019-00220-x.

[2] S. V. Bharathi, "Prioritizing and Ranking the Big Data Information Security Risk Spectrum," *Glob. J. Flex. Syst. Manag.*, vol. 18, pp. 183–201, 2017, doi: 10.1007/s40171-017-0157-5.

[3] A. Di Vaio, R. Palladino, R. Hassan, and O. Escobar, "Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review," *J. Bus. Res.*, vol. 121, pp. 283–314, 2020, doi: https://doi.org/10.1016/j.jbusres.2020.08.019.

[4] "Horizon 2020: EU framework programme for research and innovation," *Int. J. Disaster Resil. Built Environ.*, vol. 5, no. 2, Jan. 2014, doi: 10.1108/IJDRBE-03-2014-0023.

[5] M. H. Suwito, S. Matsumoto, J. Kawamoto, D. Gollmann, and K. Sakurai, "An analysis of IT assessment security maturity in higher education institution," *Lect. Notes Electr. Eng.*, vol. 376, pp. 701–713, 2016, doi: 10.1007/978-981-10-0557-2_69.

[6] A. B. Setiawan, Kautsarina, O. Rafizan, and A. S. Sastrosubroto, "Development of the information and communication technology service industry in Indonesia," *Aust. J. Telecommun. Digit. Econ.*, vol. 5, no. 3, pp. 50–82, 2017, doi: 10.18080/ajtde.v5n3.96.

[7] K.-H. Yu, A. L. Beam, and I. S. Kohane, "Artificial intelligence in healthcare," *Nat. Biomed. Eng.*, vol. 2, no. 10, pp. 719–731, 2018, doi: 10.1038/s41551-018-0305-z.

[8] T. Davenport and R. Kalakota, "The potential for artificial intelligence in healthcare," *Futur. Healthc. J.*, vol. 6, no. 2, pp. 94–98, Jun. 2019, doi: 10.7861/futurehosp.6-2-94.

[9] R. Y. Zhong, S. T. Newman, G. Q. Huang, and S. Lan, "Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives," *Comput. Ind. Eng.*, vol. 101, pp. 572–591, 2016, doi: https://doi.org/10.1016/j.cie.2016.07.013.

[10] B. Li, B. Hou, W. Yu, X. Lu, and C. Yang, "Applications of artificial intelligence in intelligent manufacturing: a review," *Front. Inf. Technol. Electron. Eng.*, vol. 18, no. 1, pp. 86–96, 2017, doi: 10.1631/FITEE.1601885.

[11] X. Yao, J. Zhou, J. Zhang, and C. R. Boër, "From Intelligent Manufacturing to Smart Manufacturing for Industry 4.0 Driven by Next Generation Artificial Intelligence and Further On," in *2017 5th International Conference on Enterprise Systems (ES)*, 2017, pp. 311–318, doi: 10.1109/ES.2017.58.

[12] S. Cockcroft and M. Russell, "Big Data Opportunities for Accounting and Finance Practice and Research," *Aust. Account. Rev.*, vol. 28, no. 3, pp. 323–333, 2018, doi: https://doi.org/10.1111/auar.12218.

[13] A. Amado, P. Cortez, P. Rita, and S. Moro, "Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis," *Eur. Res. Manag. Bus. Econ.*, vol. 24, no. 1, pp. 1–7, 2018, doi: https://doi.org/10.1016/j.iedeen.2017.06.002.

[14] Z. Rajnai and I. Kocsis, "Labor market risks of industry 4.0, digitization, robots and AI," in *2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY)*, 2017, pp. 343–346, doi: 10.1109/SISY.2017.8080580.

[15] M. Patil and M. Suresh, "Modelling the Enablers of Workforce Agility in IoT Projects: A TISM Approach," *Glob. J. Flex. Syst. Manag.*, vol. 20, 2019, doi: 10.1007/s40171-019-00208-7.

[16] J. Beilby, "Workforce innovation: Embracing emerging technologies.," *Aust. J. Gen. Pract.*, vol. 47, no. 8, pp. 522–524, Aug. 2018, doi: 10.31128/AJGP-02-18-4489.

[17] E. Colombo, F. Mercorio, and M. Mezzanzanica, "AI meets labor market: Exploring the link between automation and skills," *Inf. Econ. Policy*, vol. 47, pp. 27–37, 2019, doi: https://doi.org/10.1016/j.infoecopol.2019.05.003.

[18] A. Purnomo, T. Susanti, H. U. Anisah, A. K. Sari, and F. I. Maulana, "Value of m-commerce research: A bibliometric perspective," *Proceedings of 2021 International Conference on Information Management and Technology, ICIMTech 2021*. pp. 813–818, 2021, doi: 10.1109/ICIMTech53080.2021.9534928.

[19] T. F. and Z. M., "Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing," *IEEE Access*, vol. 5, pp. 20418 – 20427, 2017, doi: 10.1109/ACCESS.2017.2756069.

[20] W. J. *et al.*, "Brave new world: service robots in the frontline," *J. Serv. Manag.*, vol. 29, no. 5, pp. 907 – 931, 2018, doi: 10.1108/JOSM-04-2018-0119.

[21] W. Y., C. Q., H. T., and K. C., "Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3125 – 3148, 2019, doi: 10.1109/TSG.2018.2818167.