INTI JOURNAL | Vol.2025, Issue 1, No.1
eISSN:2600-7320

Tracing the Path from Industry 4.0 to Industry 5.0 through Topic Modeling Analysis

Cheng Zeng [1,2*], Hong Cheng Ding[3], Wai Yie Leong[3], Ya Fei Li[4,5]

[1]College of Physics and Electronic Information Engineering, Minjiang University, Fuzhou, China
[2]Faculty of Education and Liberal Arts, INTI International University, Nilai, Negeri Sembilan, Malaysia
[3]Faculty of Engineering and Quantity Surveying, INTI International University, Nilai, Negeri Sembilan, Malaysia
[4]College of Foreign Languages, Minjiang University, Fuzhou, China
[5]International Education College, New Era University College, Kajang, Selangor, Malaysia

[*]**Email**: i24025613@student.newinti.edu.my

Abstract

This study explores the evolution of research from Industry 4.0 to Industry 5.0 using Latent Dirichlet Allocation (LDA) to uncover key research topics and trends. This paper utilizes the Web of Science (WOS) database to collect literature on Industry 4.0 and 5.0 from 2015 to 2024. Through an LDA-based analysis, five key topics were identified, including IoT and automation, adoption frameworks, digital business transformation, smart manufacturing systems, and AI-driven models. The research highlights the growing importance of human-machine collaboration, blockchain security, and sustainable practices in the transition from Industry 4.0 to Industry 5.0. This study contributes to the understanding of evolving industrial research and offers insights into the future direction of industrial innovation.

Keywords

Industry 4.0, Industry 5.0, LDA model, Topic evolution, Research hotspots

Introduction

The notion of "Industry 5.0: Towards a Sustainable, Human-centric, and Resilient European Industry" was presented by the European Union (EU) in April 2021 (Breque et al., 2021). Industry 5.0 builds upon the principles of Industry 4.0, which emphasizes the incorporation of sophisticated technologies like IoT, AI, and cyber-physical systems into industrial processes, by prioritizing the human element in industrial systems beyond mere automation and optimization. In contrast to its predecessor, Industry 5.0 transcends mere efficiency and productivity, promoting a comprehensive

Submission: 13 September 2024; **Acceptance:** 6 November 2024; **Available online:** January 2025

strategy that emphasizes sustainability, resilience, and human welfare. This paradigm integrates technology progress with societal requirements, facilitating industries to become more inclusive and environmentally sustainable. The shift from Industry 4.0 to Industry 5.0 represents a fundamental change in industrial processes, evolving from technology-driven automation to a human-centric approach that promotes collaboration between humans and intelligent machines. Industry 5.0 reinterprets industrial innovation by incorporating sustainability and social values into the foundation of economic progress. This shift requires a comprehensive grasp of the connections between Industry 4.0 and Industry 5.0, particularly regarding their technological bases and philosophical principles. This study uses the Latent Dirichlet Allocation (LDA) model to discern research trends and topic progression in the literature concerning Industry 4.0 and 5.0. This study systematically examines research papers, elucidating the shift from technological innovation to human-centered industrial systems and providing critical insights for the design of future industries and sustainable practices.

Methodology

The main objective of this study is to collect published literature in this field, pre-process the data, and perform topic extraction to identify research hot-spots and keywords. Additionally, the paper analyzes the evolution of topics by calculating the similarity between topics at different stages. Figure 1 illustrates the comprehensive research framework, which consists of five primary stages: data collection, text preprocessing, overall LDA, DTM, and data analysis.

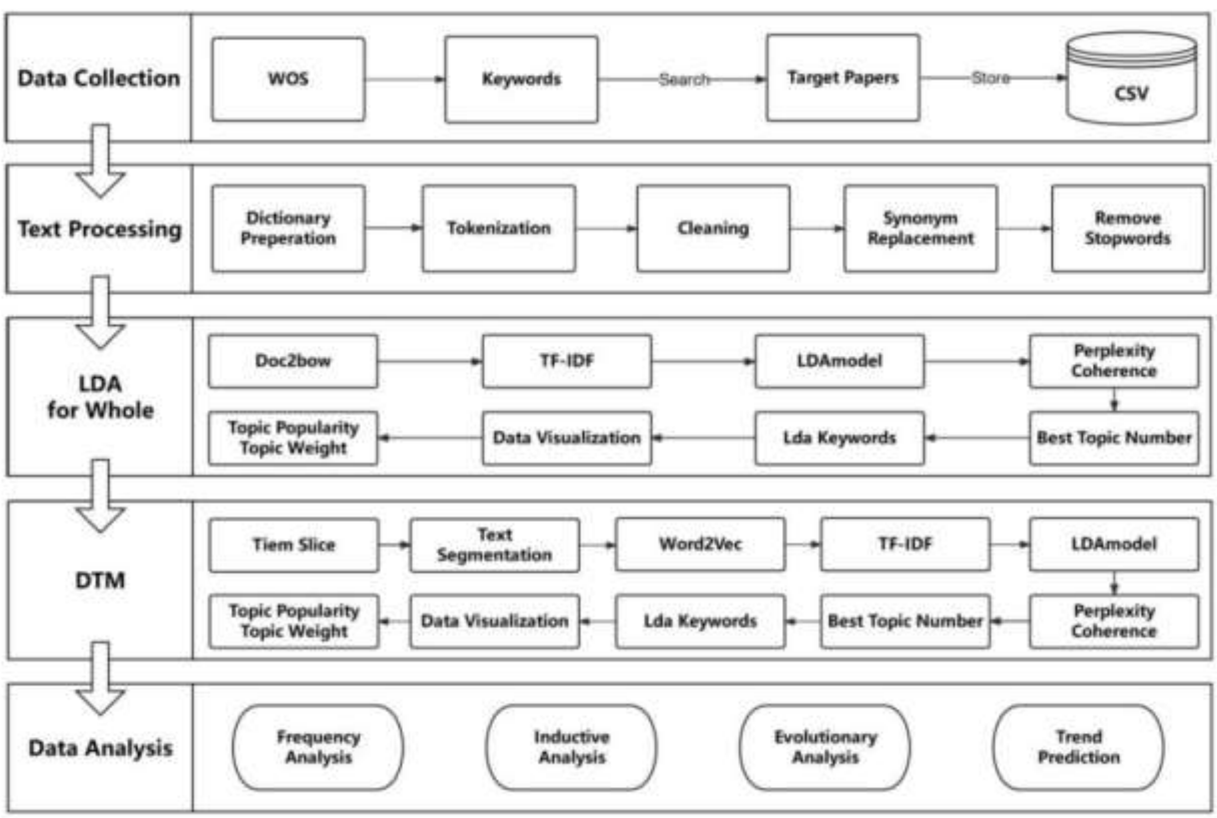

Figure 1. Flowchart of Thematic Evolution Analysis

This study employs topic searches and utilizes the WOS as the data source to gather literature data published in the fields of Industry 4.0 and 5.0 between 2015 and 2024. The literature search was performed on August 17, 2024. The search phrases used in the WOS database were TS=("Industry 4.0") and TS=("Industry 5.0") respectively. The study incorporated journal articles while excluding books, patents, and standards, and eliminating documents without abstracts. After implementing these criteria, the resulting datasets consisted of 10,835 articles related to Industry 4.0 and 1,006 articles related to Industry 5.0 written in English.

The Natural Language Toolkit (NLTK) module in Python was utilized for tokenization in this research. Next, the text underwent a series of cleaning steps including converting all words to lowercase, deleting punctuation, eliminating stop words, completing lemmatization, and omitting words that finish with 's. Afterward, the Gensim module was employed to create bigrams from the preprocessed text, and the stop word list was extended depending on the tokenization outcomes. This process effectively minimized the influence of noise in the text on the model's outcomes.

Results and Discussions

Topic Modeling Analysis

This study employs the Gensim module in Python for empirical research. The quality of the LDA model is strongly influenced by the number of potential topics (Zhao et al., 2021). In this study, coherence indicators are used to determine the optimal number of topics. Potential topic numbers were selected from the range of integers between 1 and 20, and their related coherence was assessed. Figure 2 depicts the correlation between the overall number of topics and coherence. Figure 2 displays that topic numbers 5, 7, 8, 9, and 13 possess the relatively high coherence values. However, when there are 7, 8, 9 and 13 topics, the words that make up the topics appear at the same time in several topics, which results in over-fitting. Thus, after thorough evaluation, this study has chosen an overall number of topics, specifically k = 5.

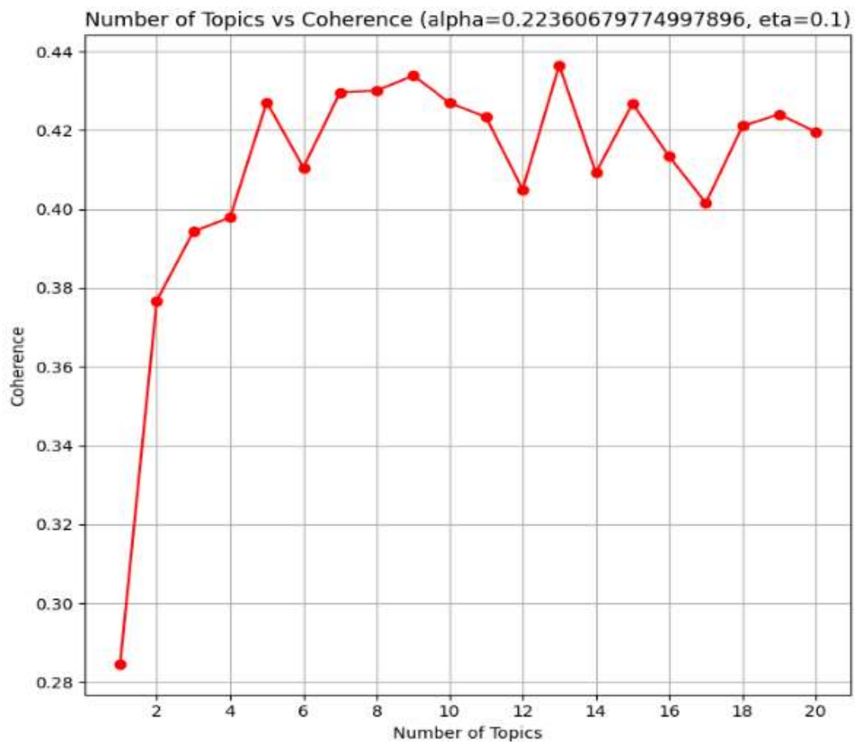

Figure 2. Coherence-Number of Topics Line Chart

Table 1 in this study displays the top 15 words with the highest frequency for each of the 5 topics. Combining the authors' domain knowledge, based on Table 1, it can be inferred that Topic 1 focuses on IoT, network systems, algorithms, and sensor-based applications, emphasizing safety and real-time control mechanisms; Topic 2 addresses the challenges and strategies of adopting Industry 4.0 and 5.0 frameworks, with an emphasis on sustainability, supply chains, and integration practices; Topic 3 covers digital transformation in business contexts, focusing on organizational management, innovation, and the influence of digitalization on business processes; Topic 4 discusses smart manufacturing systems, including process optimization, digital twins, and quality management, with an emphasis on engineering and operational effectiveness; Topic 5 explores AI-driven models and machine learning techniques for performance evaluation, emphasizing data accuracy, training, and application in various domains, including healthcare.

Table 1. Topics and frequent words

No.	Topic Name	Frequent Words	Proportion
1	Industrial IoT and Automation	network, machine, application, system, safety, control, algorithm solution, IoT, based, time, sensor, communication, environment, service	24%
2	Industry 4.0 and 5.0 Adoption Framework	technology, industry, adoption, industry_4_0, framework, challenge, supply_chain, industry_5, strategy, sustainability, benefit, environment, barrier, practice, performance	32.1%

3	Digital Transformation in Business	firm, digital, development, business, organization, influence, management, experience, innovation, technology, skill, transformation, industry future, function	26.6%
4	Smart Manufacturing Systems	manufacturing, system, process, smart, design, product, tool, part, operation, operator, concept, plant, digital_twin, quality, based	43.2%
5	AI and Data-Driven Models	model, data, based, technique, quality, application, performance, energy, AI, accuracy, feature, training, set, information, area	29.1%

Word Cloud Analysis

The word cloud for Industry 4.0 highlights terms like "technology," "system," "manufacturing," "process," and "data." This suggests a focus on the technological and systemic changes brought by Industry 4.0, particularly in manufacturing. The prominence of words like "performance," "network," "adoption," and "framework" indicates that discussions around Industry 4.0 heavily emphasize the adoption of new technologies, the creation of frameworks for implementation, and performance improvements within industrial settings.

The Industry 5.0 word cloud features terms such as "human," "system," "technology," "manufacturing," and "environment," reflecting a shift towards human-centric and sustainable approaches in industrial processes. Compared to Industry 4.0, Industry 5.0 emphasizes the integration of human elements with advanced technologies, as evidenced by the appearance of words like "human-centric," "collaboration," and "sustainability." This word cloud suggests a focus on how technology can be leveraged to not only improve efficiency but also enhance the human experience and ensure environmental sustainability.

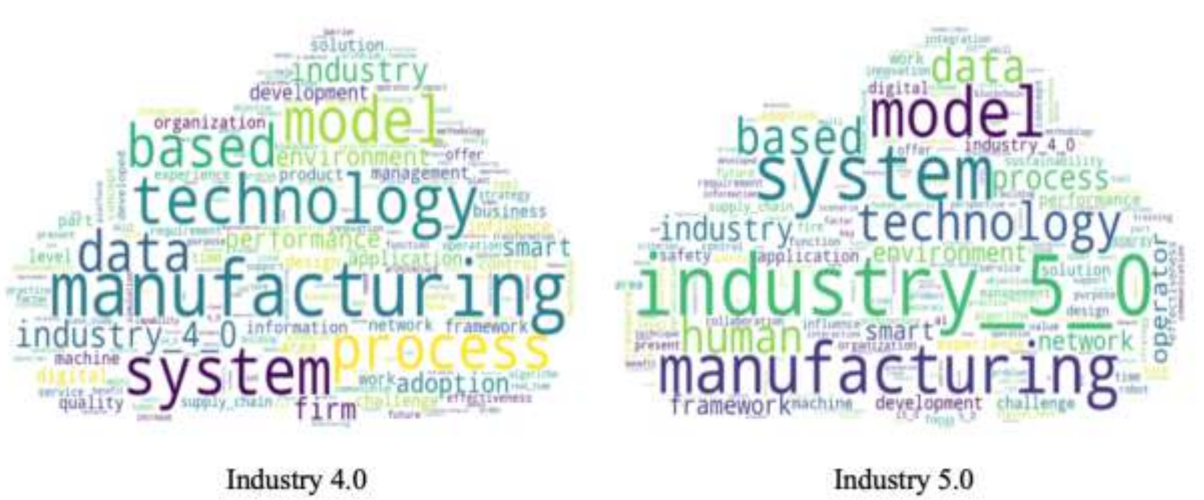

Figure 3. Word cloud by topics

Topic Evolution Analysis

Industry 4.0, which emerged as a concept in the early 2010s, emphasizes automation, interconnected systems, and the application of intelligent technologies such as the IoT, CPS, and AI (Leong et al., 2020). The research topics during this phase revolve around improving manufacturing processes through automation and data-driven technologies. For example, the "Industry 4.0 in Manufacturing and Process Quality" topic depicted in the diagram highlights the role of digital transformation in improving the quality and efficiency of manufacturing systems. Through the integration of smart systems, manufacturers aim to create intelligent, adaptable, and data-rich environments where machines communicate with one another and make autonomous decisions based on real-time data. The deployment of dynamic, model-based algorithms for smart factories, as indicated in the chart, demonstrates the application of CPS in ensuring that production systems are not only efficient but also capable of predictive maintenance and real-time optimization. These approaches are essential for improving productivity and reducing downtime in complex manufacturing environments.

AI-based prediction models and smart digital twin systems are other key Industry 4.0 technologies that represent a move toward predictive and proactive production. Digital twins, virtual replicas of physical systems, allow manufacturers to simulate, monitor, and optimize processes in real time, creating a more dynamic and flexible manufacturing environment (Huang et al., 2021). These concepts enable the transition from reactive to proactive maintenance, where systems are monitored, and potential failures are predicted before they occur. This proactive approach is further supported by data-driven machine maintenance solutions, which integrate machine learning and IoT to provide predictive insights.

However, despite the technological advancements promised by Industry 4.0, significant challenges remain, particularly concerning the integration of these technologies into existing production environments. The "Challenges of Industry 4.0 Adoption" topic in the diagram highlights the difficulties that manufacturers face when attempting to implement such complex digital systems. These challenges include the high costs of implementation, the need to upskill the workforce, cybersecurity concerns, and the interoperability of various digital systems. While Industry 4.0 largely focuses on improving manufacturing efficiency through digitalization, Industry 5.0 introduces a human-centric approach, where technology is seen as a tool for enhancing human creativity, well-being, and collaboration (Leong et al., 2023). This shift is evident in the diagram, which shows the flow from the core themes of Industry 4.0 to those of Industry 5.0. Industry 5.0 focuses on the collaborative work between humans and robots, often referred to as human-robot collaboration, as seen in the topic "Human-Robot Collaboration in Manufacturing".

In contrast to the fully automated environments envisioned in Industry 4.0, Industry 5.0 advocates for a symbiotic relationship between humans and machines, where robots act as collaborators rather than replacements for human workers. These collaborative robots, or cobots, are designed to work alongside humans, handling repetitive, dangerous, or physically demanding tasks while allowing humans to focus on more creative, complex, and decision-oriented activities. This shift is essential for ensuring that technology enhances, rather than replaces, human labor, leading to more satisfying and safer work environments. Cobots are equipped with advanced sensors and AI that allow them to adapt to human movements and operate safely in shared spaces, making them integral to the vision of Industry 5.0.

The topics related to blockchain for secure smart cities and the Industrial Internet of Things (IIoT) highlight another crucial element of Industry 5.0: the focus on resilience, sustainability, and security. In Industry 5.0, these technologies are applied to create secure, sustainable, and human-centered industrial ecosystems. Blockchain, for example, is used to ensure the security and transparency of data exchanged between interconnected systems, particularly in smart cities and IoT frameworks. This application of blockchain technology ensures that data is protected from cyber threats while maintaining the integrity and transparency of industrial processes. The emphasis on sustainability and resilience is a direct response to the limitations of Industry 4.0, which, while focused on efficiency, often neglected the broader societal impacts of automation, such as job displacement and environmental sustainability.

Human-Centric Industrial Revolution and the Evolution of Collaboration The shift from Industry 4.0 to Industry 5.0 is also marked by a broader transformation in the objectives of industrial systems. Whereas Industry 4.0 was driven by the goal of maximizing efficiency and productivity through digital and automated technologies, Industry 5.0 aims to integrate these technologies with human needs and values. This transition is captured in the "Industry 5.0: Human-Centric Industrial Revolution" theme. Here, the focus is on creating industrial systems that prioritize human well-being, creativity, and sustainability. This evolution is reflected in the increasing prominence of digital twin systems and AI-based models, which not only optimize production processes but also enhance human decision-making and creativity.

Moreover, the role of IIoT and data-driven maintenance solutions continues to play a vital role in the evolution toward Industry 5.0, but with an added emphasis on human-machine collaboration. In this context, IIoT systems are not only used to collect and analyze data from machines but also to provide humans with the insights they need to make more informed decisions. This integration of human intelligence and machine learning represents a key component of the Industry 5.0 vision.

In conclusion, the thematic flow from Industry 4.0 to Industry 5.0 reflects a significant shift in the objectives and priorities of industrial research and practice. This evolution is characterized by the integration of cobots, blockchain technologies for secure and sustainable industrial systems, and a renewed focus on resilience and human-centric innovation. The challenges of adopting these new technologies, particularly concerning safety, collaboration, and the balance between human and machine roles, remain central to the ongoing research in both Industry 4.0 and Industry 5.0.

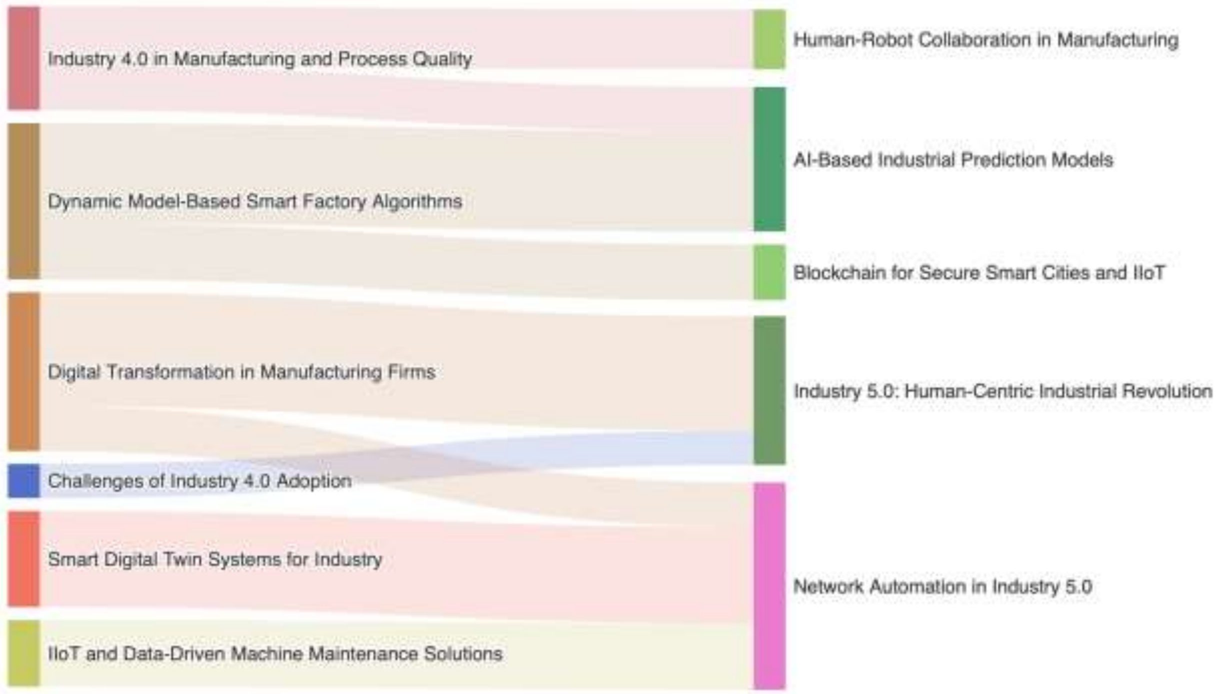

Figure 6. Topic Evolution Sankey Diagram

Conclusion

The comprehensive analysis of research trends in the field of Industry 4.0 and 5.0 reveals significant advancements and transformative impacts over the past decade. Utilizing the LDA model, the study identified five main topics: Industrial IoT and Automation, Industry 4.0 and 5.0 Adoption Framework, Digital Transformation in Business, Smart Manufacturing Systems, AI, and Data-Driven Models. In addition, the study explores the evolution from Industry 4.0 to Industry 5.0, which is characterized by a shift from pure technological optimization and automation to a more human-centered approach. While Industry 4.0 was primarily focused on increasing efficiency, productivity, and automation through advanced digital technologies, Industry 5.0 builds upon these innovations by integrating human creativity, collaboration, and sustainability into industrial systems. Research topics reflect this transition as they evolve from technical optimization (e.g., dynamic algorithms, digital twins) to human- and society-centered topics (e.g., human-robot collaboration, secure IoT frameworks, and sustainable practices). The analysis provides a foundation for understanding the current state of Industry 4.0 and 5.0, and then offers insights into future research directions.

Acknowledgement

The researcher did not receive any funding for this study, and the results have not been published in any other sources.

References

1. Breque, M., De Nul, L., & Petridis, A. (2021). Industry 5.0: Towards a sustainable, human-centric and resilient European industry. European Commission, Directorate-General for Research and Innovation. https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/industry-50-towards-more-sustainable-resilient-and-human-centric-industry-2021-01-07_en
2. Huang, Z., Shen, Y., Li, J., Fey, M., & Brecher, C. (2021). A survey on AI-driven digital twins in Industry 4.0: Smart manufacturing and advanced robotics. Sensors, 21(19), 6340. https://doi.org/10.3390/s21196340
3. Kumari, M., et al. (2021). LDA-based topic modeling for COVID-19-related sports research trends. Frontiers in Sports Research. https://doi.org/10.3389/fpsyg.2022.1033872
4. Leogn, W. Y., Leong, Y. Z., & Leong, W. S. (2023). Human-machine interaction in biomedical manufacturing. In 2023 IEEE 5th Eurasia Conference on IOT, Communication and Engineering (ECICE) (pp. 939–944). https://doi.org/10.1109/ECICE59523.2023.10383070
5. Leong, W. Y., Chuah, J. H., & Tuan, T. B. (Eds.). (2020). The nine pillars of technologies for Industry 4.0. Institution of Engineering and Technology. https://doi.org/10.1049/PBTE088E
6. Zhao, W. X., et al. (2021). Topic modeling for research trend analysis in industry sectors. Journal of Supercomputing.