Unveiling the Evolution of Enterprise Digital Innovation Strategies: Insights From U.S.-Listed Companies' Annual Reports

Shizhen Bai [ID], Yongbo Tan [ID], Chunjia Han [ID], Mu Yang [ID], Brij B. Gupta [ID], *Senior Member, IEEE*, Varsha Arya [ID], and Neeraj Kumar [ID], *Senior Member, IEEE*

Abstract—This article introduces a new metric for evaluating digital innovation in enterprise transformation using textual analysis of annual reports from U.S.-listed companies. Through network analysis and topic modeling, we identified 12 topics categorized into three main areas: digital technology innovation, customer-oriented digital strategy, and digital transformation in traditional business operations. Our research indicates that digital innovation strategies are critical for maintaining competitiveness and have shifted to a more innovation management-oriented approach. We also found differences in digital innovation strategies between companies and industries. Our study contributes to the theoretical significance of enterprise management and sustainable development.

Index Terms—Digital transformation, innovation management, social implications of technology, technology social factors, text analysis.

I. INTRODUCTION

DIGITAL transformation has become a critical priority for businesses in recent years. It refers to the integration of digital technology into all areas of a business, fundamentally changing how companies operate and deliver value to customers

Shizhen Bai and Yongbo Tan are with the School of Management, Harbin University of Commerce, Harbin 150028, P.R. China (e-mail: baishzh1962@126.com; tanyongbo1999@163.com).

Chunjia Han and Mu Yang are with the Department of Management, Birkbeck, University of London, WC1E 7HX London, U.K. (e-mail: chunjia.han@bbk.ac.uk; m.yang@bbk.ac.uk).

Brij B. Gupta is with the Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan, also with Kyung Hee University, Seoul 02447, South Korea, also with the Symbiosis Centre for Information Technology, Symbiosis International University, Pune, 412115, India, also with the Department of Electrical and Computer Engineering, Lebanese American University, Beirut 1102, Lebanon, and also with the Center for Interdisciplinary Research, University of Petroleum and Energy Studies, Dehradun 248007, India (e-mail: gupta.brij@gmail.com).

Varsha Arya is with International Center for AI and Cyber Security Research and Innovations, Asia University, Taichung 413, Taiwan, and also with Center for Interdisciplinary Research, University of Petroleum and Energy Studies, Dehradun 248007, India (e-mail: varshaarya21@gmail.com).

Neeraj Kumar is with Thapar University, Patiala 147004, India (e-mail: neeraj.kumar@thapar.edu).

Digital Object Identifier 10.1109/TCSS.2024.3360040

[1]. Social innovation, on the other hand, refers to the creation and implementation of new ideas and strategies that address social and environmental challenges [2]. There is a strong connection between digital transformation and social innovation, as digital technologies offer new opportunities for businesses to innovate and create social value. For example, digital platforms can facilitate greater collaboration and knowledge sharing among stakeholders, leading to more efficient and effective solutions to social challenges. The global progress toward digital transformation has been significant, with many companies investing heavily in digital technologies to remain competitive in the global market. The Coronavirus disease 2019 (COVID 19) pandemic has also accelerated the pace of digital transformation as companies have had to rapidly adopt remote working and digital communication technologies.

Digital transformation has the potential to drive sustainable development by enabling businesses to become more efficient and effective while reducing their environmental impact [3]. This can lead to a more inclusive and sustainable economy, benefiting both businesses and society as a whole. Given the importance of understanding the impact of digital transformation on businesses and society, it is crucial to analyze data sources such as annual corporate reports. These reports provide valuable information on a company's digital transformation efforts and can be used to gain insights into the impact of digital innovation strategies on business performance and social value creation. With the advancement of natural language processing (NLP) technology, it is now possible to extract valuable information from large-scale text data. In the context of enterprise digital transformation and the significance of annual reports, our research aims to utilize structural topic modeling (STM) to extract information from annual reports and address the following research questions: 1) what are the digital transformation strategies and behaviors adopted by enterprises? 2) how has the selection of digital strategies evolved over time? and 3) what is the relationship between industry sectors and the choice of digital strategies?

Our research makes three significant contributions to the existing literature. First, we utilize STM to extract valuable information regarding corporate digital innovation strategies from annual reports, which expands the usefulness of this rich data source beyond the accounting field [4]. Specifically, we identify three digital strategies adopted by enterprises to achieve digital transformation goals: digital technology innovation,

customer-oriented digital strategy, and digital transformation in traditional business operations. Second, our study reveals the dynamic relationship between digital innovation strategies, going beyond the findings of previous corporate digital strategy literature. We find that in the initial stages of digital transformation implementation, emphasis is placed on technological innovation, while nontechnical strategies receive more attention over time. Third, our research analyzes a broad range of industries from advanced U.S. companies, providing high-quality digital transformation templates for other countries. This contrasts with previous research that has largely focused on developed countries and highly polluting industries.

The remaining sections of this article are structured as follows. Section II presents a comprehensive literature review of existing research on social innovation and corporate digital transformation. This section provides a theoretical foundation for our research and helps establish the importance and relevance of our study. Section III outlines the methods we employed to process the text data extracted from annual reports. This includes an overview of STM and the specific steps we took to analyze the data. Section IV presents the results of our study and provides detailed insights into the digital innovation strategies adopted by advanced U.S. companies. We also conduct a comprehensive analysis of the dynamic relationship between digital innovation strategies, as well as the influence of industry on the choice of digital strategy. Finally, in Section V, we conclude the article by summarizing the main contributions of our study and highlighting directions for future research.

II. LITERATURE REVIEW

Our study focuses on examining the digital transformation strategies of publicly listed companies in the United States within the context of social innovation. By analyzing the content of corporate annual reports, we aim to uncover the dynamic relationship between various digital transformation strategies employed by these companies. The significance of digital transformation and innovation cannot be overstated in today's fast-paced business environment. Companies are increasingly relying on digital technologies to improve their operations, enhance customer experience, and drive growth. Furthermore, there is a growing recognition of the importance of social innovation and sustainable development in achieving business success.

We aim to provide a comprehensive review of the existing literature on digital transformation and innovation, with a particular emphasis on their intersection with social innovation and sustainable development. We also justify our research methods by summarizing previous studies that have utilized topic models for information extraction from corporate-related texts.

A. Social Innovation and Sustainable Digital Transformation

Enterprises today face an increasingly complex environment that requires them to adopt new approaches to create value and achieve sustainable growth. Social innovation has emerged as a critical tool for organizations to drive digital transformation and sustainable development. By focusing on the needs of society and the environment, enterprises can create innovative solutions that generate economic and social benefits while also minimizing their environmental impact. This section examines the current research on how enterprises can leverage social innovation to promote digital transformation and sustainable development.

Numerous studies underscore social innovation's role in driving digital transformation and sustainability. Di Vaio et al. [5] conducted a literature review, affirming social innovation's vital role in addressing challenges, fostering sustainable business models, and stimulating technological advancement. Sustainable development, pivotal for enterprises balancing economic, social, and environmental concerns, benefits from social innovation's stakeholder-oriented solutions. Lüdeke-Freund [6] further asserts that social innovation empowers enterprises to adopt sustainable practices, generate value, and address pressing issues. Digital transformation integrates digital tech into operations, enhancing efficiency and addressing challenges. Schaffers et al. [7] highlight the link between social innovation and digital transformation, with social innovation fueling innovation and digital business models. Guandalini [8] conducted a systematic literature review, shedding light on the multifaceted connections between these domains, underscoring social innovation's potential in addressing societal and environmental challenges while stimulating technological advancements. Moreover, Busulwa et al. [9] explored the digital hospitality ecosystem, emphasizing the impact of digital transformation and social innovation in the tourism sector. Bocken et al. [10] proposed sustainable business model archetypes, providing a framework for organizations to integrate social innovation, digital transformation, and sustainability into their operations. Piller et al. [11] delved into the cocreation of innovation through social media, highlighting its influence on the integration of social innovation into digital processes. Miklosik and Evans [12] conducted a literature review on learning and digital innovation, shedding light on the role of learning processes in fostering social innovation within digital contexts. Westley and Antadze [13] emphasized strategies for scaling social innovation, emphasizing its potential for greater societal impact. Phills et al. [14] rediscovered social innovation, emphasizing its relevance in addressing contemporary societal challenges. These studies collectively underscore the intricate relationship between social innovation, digital transformation, and sustainable development, offering fresh perspectives on how organizations can harness the synergies among these domains.

Corporate social responsibility (CSR) is another essential factor that enterprises must consider when promoting social innovation and sustainable development. CSR involves integrating social and environmental concerns into business operations and decision-making. Cornelius et al. [15] argued that CSR can play a vital role in promoting social innovation by creating a culture of innovation that focuses on social and environmental concerns. CSR can also help enterprises to identify and address social and environmental challenges, create new partnerships with stakeholders, and leverage digital technology to create innovative solutions. Romero et al. [16] investigate the application of digital tools in CSR and sustainability reporting within Spanish universities, highlighting the role of digitalization in enhancing accessibility and dissemination of CSR and sustainability information. Xu et al. [17] adopt a dynamic capabilities approach to analyze

how digital transformation impacts CSR practices, emphasizing the potential for improved CSR performance through digitalization. Additionally, the McKinsey Global Institute's report on artificial intelligence (AI) [18] delves into the pivotal role of AI in digital transformation and discusses its influence on CSR strategies and interactions between businesses and society, offering insights into the evolving landscape of CSR in the digital era. Cumulatively, these studies underscore the multifaceted relationship between CSR, digitalization, and social innovation, highlighting the transformative potential of CSR in fostering innovation, the role of digital tools in CSR reporting, and the evolving dynamics within CSR influenced by digital transformation and AI.

Social innovation is an essential tool for enterprises to drive digital transformation and sustainable development. By focusing on the needs of society and the environment, enterprises can create innovative solutions that generate economic and social benefits while also minimizing their environmental impact [19]. This section highlights the critical role of social innovation in promoting sustainable development, the integration of digital technology, and the importance of CSR. Further research is needed to explore the specific strategies and best practices that enterprises can use to leverage social innovation to promote sustainable development and drive digital transformation.

B. Digital Transformation Strategy

Digital transformation has become a significant priority for many companies seeking to remain competitive in today's fast-paced business environment. Digital transformation involves the integration of digital technologies into all aspects of a company's operations, leading to improvements in efficiency, productivity, and customer experience. The purpose of this section is to explore innovative strategies that companies can employ to successfully undertake digital transformation.

Fostering a digital culture is the initial step in successful digital transformation. It entails embracing digital technologies, promoting experimentation, and valuing collaboration [20]. Developing digital capabilities, through training, partnerships, and talent acquisition, equips companies to seize digital opportunities and overcome transformation challenges [21]. Data and analytics are pivotal, aiding in understanding customer preferences, optimizing processes, and innovation [22]. Embracing emerging technologies like AI, IoT, and machine learning enhances competitiveness and transforms business operations [23]. Creating ecosystem partnerships with other organizations provides access to technology and expertise, driving innovation and customer value [24].

Recent research has investigated various aspects of digital transformation strategies. Pinto Filho et al. [25] conducted a systematic review to uncover insights and challenges guiding organizations through digital transformations. Avdeeva et al. [26] explored innovative digital transformation models, emphasizing cocompetition and the transformation of value chains into networks within business ecosystems. Yucel [27] highlighted the importance of robust digital transformation strategies, introducing a generic evaluation and modeling model that considers market dynamics, digital disruption characteristics, corporate attributes, and implementation processes. In the context of China's digital transformation era, Lei and Zhaolei [28] emphasized the role of technologies like information technology, mobile Internet, cloud computing, the Internet of Things, and AI. They leveraged SWOT analysis to enhance the competitive advantage of enterprises in the digital economy environment. Last, Behare et al. [29] underscored the significance of effective digital business strategies, focusing on the alignment of organizational intentions with customer expectations and achievement. These studies collectively contribute to our understanding of digital transformation strategies and their implications for organizations in the evolving digital landscape.

Digital transformation is a complex process that requires a multifaceted approach. Companies seeking to successfully undertake digital transformation must develop a digital culture, build digital capabilities, leverage data and analytics, embrace emerging technologies, and create ecosystem partnerships. By implementing these strategies, companies can improve efficiency, productivity, and customer experience, and remain competitive in today's rapidly changing business environment. Our research aims to explore the practices and strategies of enterprises using social innovation for social transformation from a new perspective of annual reports using machine learning methods.

C. Applications of the Topic Model

As NLP techniques mature, more and more academics are studying texts published by firms. The vast majority of scholars use the STM, a statistical model used to discover abstract topics in a series of documents in fields such as machine learning and NLP, and this machine learning-based approach has been used extensively in the processing of unstructured data [30]. This line of approaches provides the methodology for this study, we review and discuss its use in recent business studies as follows.

The texts published by firms contain mainly corporate annual reports, earning call transcripts, CSR reports, corporate environmental reports, analyst reports, etc. Jagannathan and Niveditha's study used topic model to extract strategic topics and CSR-related content from corporate annual reports, individually [31], [32]. Huang's research used topic modeling for corporate earning call transcripts from which information of interest to investors was extracted, in turn, to explore the role analysts play in investment behavior [33]. Bellstam developed a new measure of innovation through a topic model using the text of analyst reports from S&P 500 firms [34]. In the study of Li, a cultural dictionary was created using a STM and 209 480 earnings conference call transcripts. The five key corporate culture values of innovation, integrity, quality, respect, and teamwork were then rated [35]. Pröllochs used topic model for CSR reporting to identify and address the challenges faced by firms [36]. Bai et al. [37], [38] analyzed 112 000 reviews of Shanghai Disney Resort, finding issues impacting tourist enjoyment and variations based on ticket types and group sizes. Bai et al. [38] explored themes in 112 000 visitor reviews, highlighting service and playing experience, shifts in early and later tourists' focus, and changing review-rating relationships over time. Chae and Olson [39] examined data from 133 344 U.S. companies, identifying key digital technologies affecting new

The Flow Chart of Our Machine Learning Approach

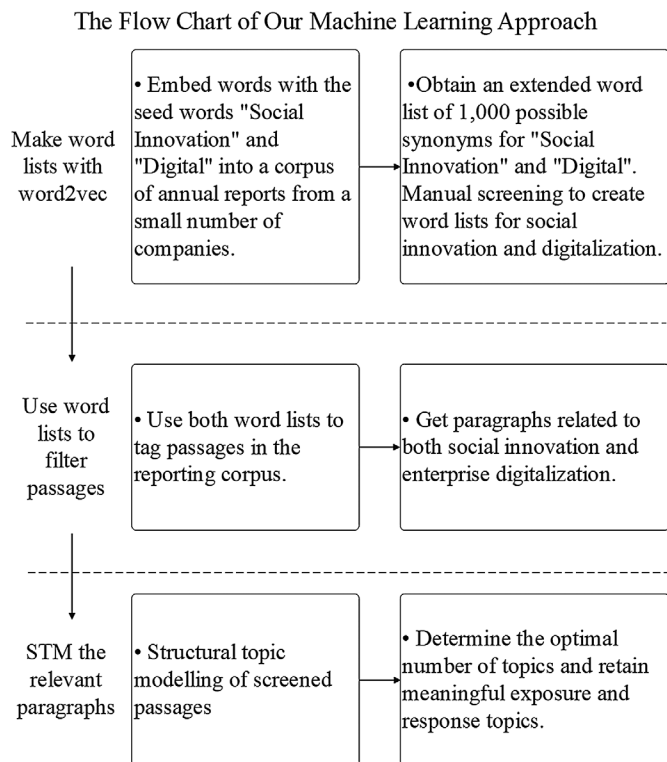| Make word lists with word2vec	• Embed words with the seed words "Social Innovation" and "Digital" into a corpus of annual reports from a small number of companies.	•Obtain an extended word list of 1,000 possible synonyms for "Social Innovation" and "Digital". Manual screening to create word lists for social innovation and digitalization.
Use word lists to filter passages	• Use both word lists to tag passages in the reporting corpus.	• Get paragraphs related to both social innovation and enterprise digitalization.
STM the relevant paragraphs	• Structural topic modelling of screened passages	• Determine the optimal number of topics and retain meaningful exposure and response topics.

Fig. 1.   Flow chart of our machine learning approach.

business creation and their industry sector associations. These studies enhance understanding of user-generated data dynamics, tourism, and digital technologies in business ventures.

To summarize, these studies employed topic models to analyze corporate-related texts and attest that the richness of the information contained in corporate-published texts can enable the extraction of insightful results.

III. METHODOLOGY

In this section, we describe the topic of our use of corporate annual reports to measure digital transformation at the company level. Fig. 1 shows a flowchart of our machine learning approach.

A. Data Acquisition and Word2vec

For this study, we obtained disclosed annual reports of companies from the official website of the U.S. stock exchange (available at https://www.sec.gov/edgar/searchedgar/company-search). We selected the annual reports of listed companies in the United States as research data for several reasons. The United States has one of the largest economies in the world, with a vast number of listed companies that are widely distributed, providing a rich source of research material. Furthermore, the United States has advanced experience and practices in corporate social innovation and digital transformation, making the study of annual reports of listed companies in the U.S. valuable for experience and inspiration. The financial and securities trading markets in the United States are relatively mature and stable, and listed companies must comply with strict and standardized disclosure requirements for their financial and business data.

Consequently, these reports provide a reliable data base for research. Additionally, the financial and business data of listed companies in the United States are relatively transparent, providing a reliable data base for research.

As an annual report includes irrelevant information such as company operating conditions and emission reduction data, we needed to filter out paragraphs related to the digitalization and innovation of the company. The word embedding model is based on a simple, time-tested concept in linguistics: Words that cooccur with the same neighboring words have similar meanings [40]. The model thus converts the neighboring word counts of a word to a numerical vector, which captures the meaning of the word and supports a synonym search using vector arithmetic. While there are different variants of the word embedding model, we use a popular neural network model, word2vec [41], to efficiently learn dense and low-dimensional word vectors. In essence, word2vec "learns" the meaning of a specific word via a neural network that "reads" through the textual documents and thereby learns to predict all its neighboring words. The output from the process is a vector representation of the word once learning has been completed after a number of iterations through the documents. The vector has a fixed dimension and captures the properties of the original cooccurrence relationship between the word and its neighbors.

We use the *gensim* library in Python to train the word2vec model. We set the dimensionality of the word vector to 300, define two words in a sentence that are no more than five words apart as neighboring words, and ignore words that occur less than five times in the corpus. After training, the model converts each word in the call corpus into a 300-dimensional vector representing the meaning of that word; we can then calculate the cosine similarity between any two word vectors to quantify the association between them. Using this feature, we construct a social innovation word list by associating a set of words collected from annual reports with social innovations. We then select the top 100 words that are most closely associated with the social innovation word vectors (i.e., have the highest cosine similarity between word vectors). We do not consider the named entities that are automatically recognized by the *CoreNLP* package. We manually checked all words in the automatically generated list and excluded those that did not meet the requirements. Most of the excluded words were either too general in meaning or too specific in terms of industry context. With the word lists for social innovations, we constructed digitized word lists using a unified methodology, tagging passages in which any of the words from the two lists appeared at the same time.

After paragraph filtering, we used NLP to preprocess the documents to make them compatible with unsupervised text mining technology. Using Python, we cut the Chinese text, saved specific nouns, and deleted certain stop words. After preprocessing, we obtained a corpus of 5555 processed paragraphs and 491 877 unique words, or vocabulary, from the annual reports.

B. Illustration of STM

In order to explore firm-level strategies and actions related to digital innovation, we need to identify discussion topics related to digitization. We use the STM [42] to accomplish this task.

TABLE I
COMPARISON OF LDA AND STM

LDA	STM
Covariates are not allowed to be included in document modeling	Information other than textual content such as time, some characteristic of the sample, etc. can be added to the probabilistic topic modeling process
No correlation between topics allowed	Allows for the existence of themes interconnectivity between topics and allows estimation of topic-to-topic relationships
Indicators for selecting high-frequency words are relatively homogeneous	High-frequency terms under various metrics were considered, such as Probability, FREX, Lift and Score.

STM is an innovative improvement on Latent Dirichlet allocation (LDA), a more elementary topic modeling approach pioneered by Blei et al. [43]. In recent years, there has been a gradual increase in the use of topic modeling in quantifying the textual disclosures of firms, such as earnings calls [33]. The principle of LDA is a process that mimics the composition of a document (e.g., an environmental report) by humans. Specifically, LDA assumes that each word in a document is generated through the following two steps. First, assuming that the content of a human-clutched document m is a specific set of topics and can be described by a distribution θm, one of the topics is drawn according to this distribution. Next, assuming that the drawn topic k has its own lexical distribution βk, a lexicon is similarly drawn according to this distribution. Each vocabulary generation goes through these two steps until this document is fully generated. LDA's inference algorithm iteratively discovers the topic distribution for each document and the word distribution for each topic by fitting this two-step generative model to the words observed in the document collection (i.e., the corpus) until the best set of parameters describing the topic and word distributions is found. The fitted model provides the following: 1) topic proportions, which tell us the weight of each topic in a document, and 2) word distributions, which provide a list of the most likely words associated with a given topic.

STM is similar to LDA, the difference being that STM will give connections between topics. Thus, STM is a more realistic generative model than LDA, providing a better model fit [44]. The advantages and disadvantages of LDA and STM are shown in Table I. Its advantage is that it allows merging metadata of data samples to explain local popularity [42], [45]. Conceptually, the interpretation of the meaningful parameters estimated by STM is almost identical to that of LDA. We can decompose a document into a mixture of topics whose proportions sum to one, and we can also label these topics by examining the distribution of words for each topic. We used the *stm* package in R to fit the STM according to the variance expectation maximization algorithm developed by Roberts et al. [46], in which the content of annual reports and unique words are input as documents and vocabulary words, respectively. Fig. 2 shows the plate of STM and visualizes the text analysis process.

Next, we need to determine the number of topics k, which is an important user-specified parameter of STM and is helpful for substantive interpretation of modeling results [48].
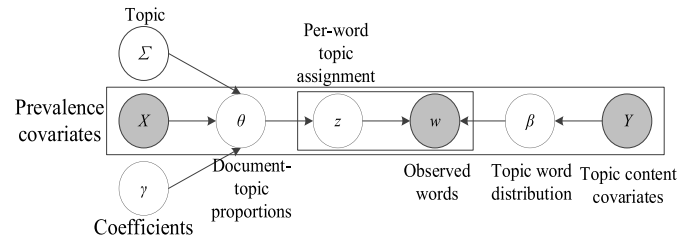

Fig. 2. Graphical illustration of STM approach [47].

C. Topic Number Estimation

After we input the data, the STM model outputs a list of words for each topic that have the highest probability of occurring in that topic. Two important criteria for judging the explanatory power of the model output are semantic coherence and exclusivity. Although we extracted these passages from the articles that contained the specified topic words, they required a certain degree of coherence to make the output more meaningful, and the semantic fragmentation caused the actual semantics of the topics to be incorrectly captured. The higher the score for semantic coherence, the more frequently the most commonly used words in a topic appear together. In addition, exclusivity is used to measure the extent to which a topic is distinguished from other topics and is based on a comparison of the similarity of word distributions. The higher the exclusivity score, the less likely it is that a high-frequency topic word in this topic will appear in other subjects.

To explore the inventive measures and approaches that organizations have employed to accomplish their digital transformation objectives, we have included the year of publication of each company's annual report and the company's stock code as metadata variables at the document level. This information will help provide a comprehensive understanding of the digital transformation strategies and tactics adopted by various companies over time. By using the functions *searchK* in the *stm* and *furrr* packages in R, the topic number K from 5 to 20 associated with the semantic coherence of the topics, held-out likelihood and residuals are shown in Fig. 3.

Fig. 3 shows an overall decreasing trend in semantic coherence and residuals as the number of topics in the model increases, but held-out likelihood scores maintain an overall increase. After taking into account the semantic coherence of topics, the held-out likelihood and residuals, a 12-topic modeling scheme was chosen for this study.

IV. RESULTS AND ANALYSIS

In this section, we present an introduction to the relevant topics and their interconnections. Subsequently, we examine the temporal variation of the topic ratios. We then conduct an analysis of the digital disclosure practices of firms across various industries and explore the differences in the innovations and strategies that these industries have put in place for digital transformation.

A. Topic Labeling and Summary

Our study employs a statistical approach called STM to identify and classify 12 distinct topics related to corporate digital
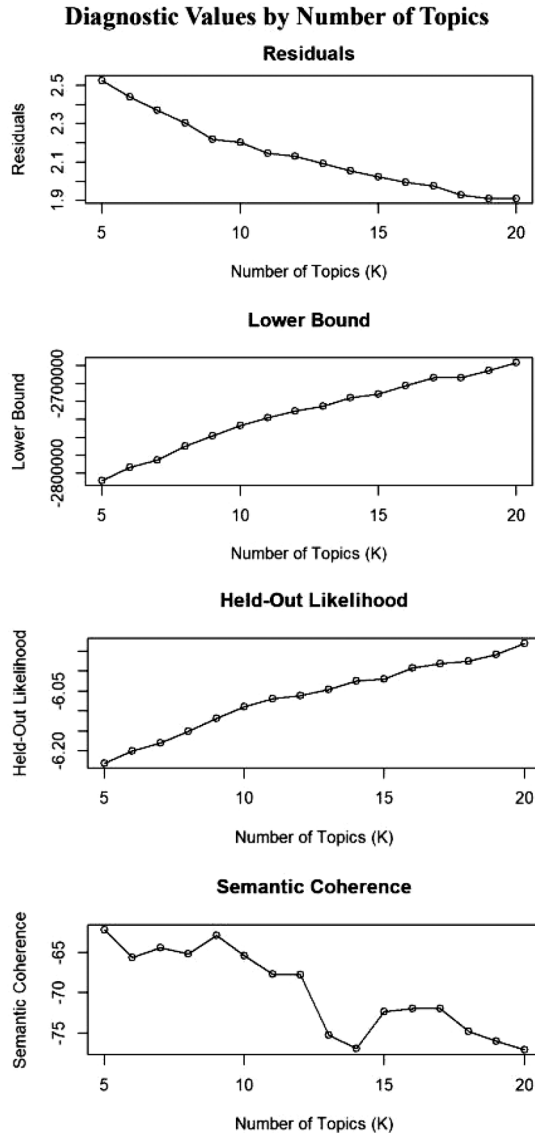
Diagnostic Values by Number of Topics

Fig. 3. Selective topic model solution.

TABLE II
TOPIC LABELING AND TOP WORDS

Topic	Topic Label	Top Word
Topic 1	Digital solutions of surveying and data analysis in various sectors	Product, system, design, manufacture, digital, technology, software, process, improvement, develop
Topic 2	Cognizant consulting and technology solutions for digital transformation	Digital, business, customer, service, technology, solution, management, transformation, operation, innovation
Topic 3	Omnichannel retail strategy for seamless customer experience	Custom, digit, retail, store, experience, continuous, channel, commerce, online, improve
Topic 4	Corporate financing through issuance of senior unsecured notes	Company, statement, report, asset, acquisition, digital, expectation, estimate, capital, cash
Topic 5	Digital risk management in business operations	Business, operations, including, risk, regulation, impact, systems, security, results, adverse
Topic 6	Digital payments in the financial industry	Payment, service, financial, card, provide, bank, include, transaction, customer, solution
Topic 7	Digital commerce sales included in comparable store sales reporting	Revenue, sales, increase, fiscal, cost, comparison, segment, expense, operation, growth
Topic 8	Omnichannel media planning and optimization tool	Market, advertise, brand, medium, consumer, digital, social, product, include, company
Topic 9	Digital competition	Product, competitive, market, technology, consumer, ability, development, business, increase, company
Topic 10	Digital entertainment content distribution	Content, service, game, digit, offering, video, entertainment, provide, online, distribute
Topic 11	Wireless communications technology and patent innovation	Data, network, security, technology, cloud, service, provide, device, storage, solution
Topic 12	Digitalization in healthcare	Data, network, security, technology, cloud, service, provide, device, storage, solution

transformation measures and strategies across a corpus of 4564 annual reports spanning from 2011 to 2022. Each topic represents a fundamental distribution of words, and each word in the dataset is assigned to one of these 12 topics. The STM method outputs a list of the most frequently occurring high-frequency words in each topic. To assign labels and categories to each topic, we recruited two social science researchers who discussed and analyzed the word lists and identified relevant examples of paragraphs with the highest percentage of vocabularies under each topic. Our research team then assigned labels and categories to each topic based on their analysis of these examples. We present these labels, along with the ratio of each topic to the top ten words, in descending probability order, in Table II. Additionally, we illustrate the proportion of these 12 topics in the total enterprise annual report in Fig. 4.

As depicted in Fig. 4, the top three topics in business operations for the annual reports of listed companies in the United States during the period of 2011–2022 were "cognizant

consulting and technology solutions for digital transformation," "digital competition," and "digital risk management in business operations." These topics received significant attention from companies during this period, possibly due to the escalating significance of digital transformation and competition in the business environment, as well as the necessity for enterprises to efficiently manage digital risks in their operations. The topic of "cognitive consulting and technical solutions for digital transformation" suggests that many companies are actively seeking external expertise and support to attain digital transformation. The topic of "digital competition" signifies that companies are focusing on strategies to sustain competitiveness in the rapidly evolving digital market. The topic of "digital risk management in business operations" highlights that companies acknowledge the need to proactively manage the risks associated with digital technology and processes in their operations. In essence, the prevalence of these topics indicates that digital transformation is
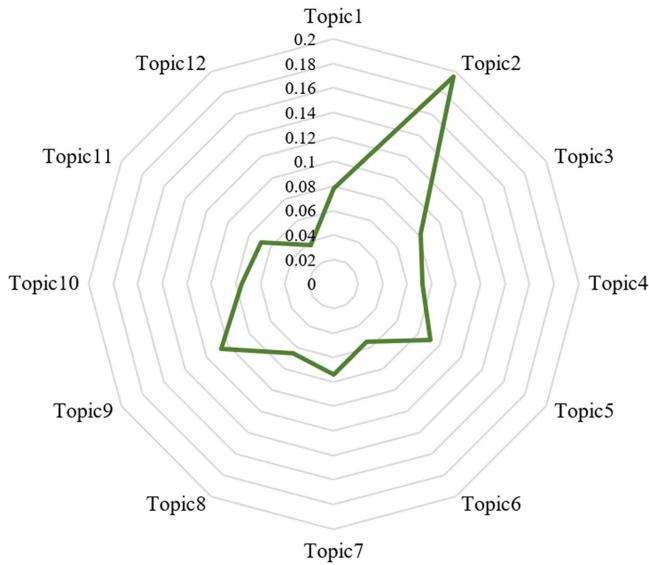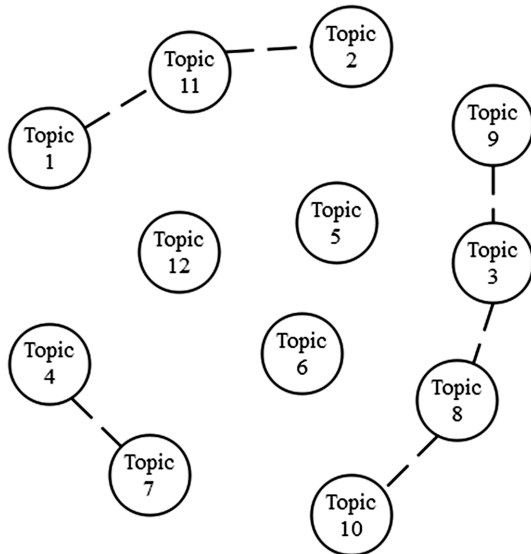
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

BAI et al.: UNVEILING THE EVOLUTION OF ENTERPRISE DIGITAL INNOVATION STRATEGIES

7

Fig. 4. Topic distribution.

Fig. 5. Topics related map.

TABLE III
TABLE OF TOPIC GROUPINGS

Category	Topic Label
Digital technology innovation	Topic 2: Cognizant consulting and technology solutions for digital transformation Topic 11: Wireless communications technology and patent innovation Topic 1: Digital solutions of surveying and data analysis in various sectors
Customer-oriented digital strategy	Topic 9: Digital competition Topic 3: Omnichannel retail strategy for seamless customer experience Topic 8: Omnichannel media planning and optimization tool Topic 10: Digital entertainment content distribution
Digital transformation in traditional business operations	Topic 4: Corporate financing through issuance of senior unsecured notes Topic 7: Digital commerce sales included in comparable store sales reporting
Uncategorized topics	
Topic 5: Digital risk management in business operations	
Topic 6: Digital payments in the financial industry	
Topic 12: Digitalization in Healthcare	

Based on the association diagram in Fig. 5, the 12 topics identified by STM were categorized into three main groups, with three topics remaining unclassified, indicating that they were less related to the other categorized topics. Through in-depth discussion and analysis, we have named the three categorized groups of topics, as presented in Table III. This approach leverages network analysis and STM package capabilities in R, which simulate the relationships between topics and categories based on word cooccurrence, providing a useful visual representation of topic associations.

1) Digital Technology Innovation: The first group comprises three topics: topic 1, "digital solutions of surveying and data analysis in various sectors"; topic 2, "cognitive consulting and technical solutions for digital transformation"; and topic 11, "wireless communication technology and patent innovation." Given their common focus on achieving digital transformation through digital technology innovation, we propose to name this group "digital technology innovation." "By combining these themes, organizations can better understand the potential benefits of developing new technologies to achieve digital transformation, including improving efficiency, improving decision-making, and taking appropriate actions to leverage these benefits."

Topic 1 highlights the importance of digital solutions for surveying and data analysis across various sectors, including construction, agriculture, and mining. Digital tools and techniques enable organizations to collect and analyze data more efficiently and accurately, leading to better decision-making and improved business outcomes. Topic 2 emphasizes the need for digital transformation in organizations and the role of consulting and technology solutions in achieving this transformation. This topic highlights the benefits of digitalization, such as increased efficiency, improved customer experience, and enhanced innovation. Organizations can leverage these benefits to gain a competitive advantage. Topic 11 focuses on wireless communications technology and patent innovation,

a crucial issue for companies, and it is wise for them to engage external support to address digital competition among enterprises and effectively tackle digital risks.

To analyze and discuss the 12 topics identified by STM, we adopted a network analysis approach, which is a commonly used method in the field of topic modeling. The STM package in R was used to simulate the relationship between topics and categories based on word cooccurrence, and a graph of associations between topics was generated. Fig. 5 displays the associations between the topics. Through this approach, we were able to uncover the interrelatedness and hierarchy among the identified topics and gain a deeper understanding of the underlying topics and concepts. The network analysis approach allowed us to explore the complex relationships between topics and provided a comprehensive visualization of the results, which can facilitate the interpretation and communication of our findings.

which play a crucial role in advancing digitalization and driving organizational performance. This topic emphasizes the importance of patent innovation in promoting technological progress and enabling organizations to stay competitive in the digital era.

One example of how organizations can leverage this group's insights is by adopting digital solutions to improve their business operations and enhance their customer experience. For instance, organizations can use digital tools to collect and analyze data in real-time, identify potential issues, and take corrective actions promptly. Additionally, they can leverage wireless communications technology to enhance their connectivity, improve their agility, and drive innovation.

2) Customer-Oriented Digital Strategy: The second group consists of topic 9 "digital competition," topic 3 "omnichannel retail strategy for seamless customer experience," topic 8 "omnichannel media planning and optimization tool," and topic 10 "digital entertainment content distribution." These topics collectively emphasize the importance of leveraging digital tools and strategies to gain competitive advantage in the market and provide a seamless and satisfying customer experience across multiple channels. Therefore, we name this group "customer-oriented digital strategy."

Topic 9 "digital competition" highlights the increasing competition in the digital space, where businesses compete to gain market share and customer loyalty. This topic emphasizes the importance of understanding digital competition dynamics and using digital tools and strategies to gain a competitive edge. Topic 3 "omnichannel retail strategy for seamless customer experience" focuses on the importance of providing a seamless customer experience across multiple channels, such as brick-and-mortar stores, online stores, and mobile apps. This topic emphasizes the benefits of an omnichannel retail strategy, such as increased customer satisfaction, loyalty, and sales. Topic 8 "omnichannel media planning and optimization tool" emphasizes the importance of using digital tools and techniques to optimize media planning across multiple channels, such as social media, email marketing, and online advertising. This topic highlights the benefits of an omnichannel approach to media planning, such as improved targeting, efficiency, and effectiveness. Topic 10 "digital entertainment content distribution" focuses on the increasing use of digital channels to distribute entertainment content, such as music, movies, and TV shows. This topic emphasizes the benefits of digital distribution, such as wider reach, greater convenience, and lower distribution costs.

3) Digital Transformation in Traditional Business Operations: The third group consists of topic 4, "corporate financing through the issuance of senior unsecured notes," and topic 7, "digital commerce sales included in comparable store sales reports."

Topic 4 "corporate financing through issuance of senior unsecured notes" highlights the use of senior unsecured notes as a financing option for corporations. This topic emphasizes the benefits of senior unsecured notes, such as flexibility, lower costs, and simpler terms, as compared to other forms of financing. The use of digital solutions in issuing and trading these notes is also gaining traction, offering greater efficiency and transparency. Topic 7 "digital commerce sales included in comparable store sales reporting" focuses on the inclusion of digital

commerce sales in comparable store sales reporting. This topic highlights the importance of accurately measuring and reporting digital commerce sales, given their increasing significance in retail operations. The use of digital solutions in tracking and analyzing these sales is becoming more prevalent, offering better insights and decision-making capabilities to retailers.

Based on these similarities, we name this group "digital transformation in traditional business operations." This group highlights the importance of leveraging digital solutions to transform traditional business practices and stay competitive in a rapidly evolving market.

Despite differences, both topics emphasize the increasing role of digital solutions in traditional business operations. Topic 4 shows how digital solutions can change the financing landscape, while topic 7 emphasizes how digital commerce can become an important component of retail operations. Together, these topics demonstrate how digitalization permeates different aspects of the business and is increasingly becoming an essential component of success. Based on these similarities, we name the group "digital transformation in traditional business operations." The group emphasized the importance of using digital solutions to transform traditional business practices and remain competitive in rapidly evolving markets.

4) Remaining Topics: The remaining three topics that are not grouped are topic 5 "digital risk management in business operations," topic 6 "digital payments in the financial industry," and topic 12 "digitalization in healthcare."

Topic 5 "digital risk management in business operations" focuses on the use of digital tools and techniques to identify, assess, and mitigate risks in business operations. This topic highlights the importance of leveraging digital solutions to identify potential risks in real-time and prevent or mitigate their impact on business operations. Topic 6 "digital payments in the financial industry" highlights the increasing use of digital payment methods in the financial industry, such as mobile payments, contactless payments, and digital wallets. This topic emphasizes the benefits of digital payments, such as convenience, security, and cost-effectiveness, and how they are transforming the financial industry. Topic 12 "digitalization in healthcare" focuses on the use of digital technologies to improve healthcare services, such as telemedicine, electronic health records, and digital diagnostic tools. This topic highlights the potential of digitalization to improve the accessibility, affordability, and quality of healthcare services, especially in remote or underserved areas.

These three topics are not directly related to the other topics in the dataset, and they highlight the increasing importance of digitalization in different sectors. By including these topics, the analysis provides a more comprehensive view of the digital transformation that is taking place in various industries.

In summary, the first group (topics 1, 2, and 11) highlights the importance of digital solutions in technology innovation. The second group (topics 3, 8, 9, and 10) emphasizes the importance of omnichannel strategies in retail and digital commerce. The third group (topics 4 and 7) focuses on corporate financing through the issuance of senior unsecured notes. Finally, the remaining three topics (topics 5, 6, and 12) highlight the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

BAI et al.: UNVEILING THE EVOLUTION OF ENTERPRISE DIGITAL INNOVATION STRATEGIES   9
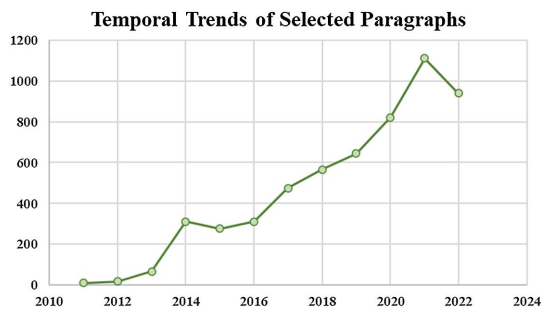

Fig. 6.   Temporal trends of paragraphs in annual reports.

increasing importance of digitalization in business operations, financial industry, and healthcare.

B. Changes in Topic Over Time

To illustrate the temporal distribution of the selected paragraphs from the annual reports of U.S.-listed companies, we plotted the number of paragraphs filtered by the innovation and digitization dictionaries for each year from 2011 to 2022, as shown in Fig. 6. The figure provides an overview of the quantity of relevant content in the annual reports over time and serves as a basis for further analysis of the evolving trends in the digital transformation and innovation strategies of the companies.

The increasing trend of the number of selected paragraphs related to digital transformation and innovation strategies in the annual reports of U.S. public companies from 2011 to 2022 can be explained by the growing importance of digital transformation for businesses to maintain a competitive edge and adapt to changing market demands. Companies have been investing in digital technologies and platforms to improve operational efficiency, enhance customer experience, and explore new business models. The COVID-19 pandemic in 2020 has further accelerated the adoption of digital transformation strategies as businesses faced disruptions in supply chains, operations, and customer interactions. Many companies have leveraged digital technologies to facilitate remote work, enable e-commerce, and enhance digital marketing efforts. As a result, the number of selected paragraphs related to digital transformation and innovation strategies in 2020 exceeded that of 2021, as companies had already implemented many digital transformation initiatives in response to the pandemic.

Overall, the increasing trend of selected paragraphs related to digital transformation and innovation strategies in U.S. public companies' annual reports suggests that digital transformation has become a critical component of companies' strategic agendas, and businesses will continue to invest in digital technologies to enhance their competitiveness and resilience in the future.

The analysis of trends in various topics is an essential aspect of understanding the developments and changes in different industries. As businesses strive to stay competitive and adapt to the changing landscape of technology and markets, it is critical to identify emerging trends and shifts in consumer behavior. In this section of the analysis, we will examine the trends in twelve topics related to digital solutions, consulting, financing, retail,
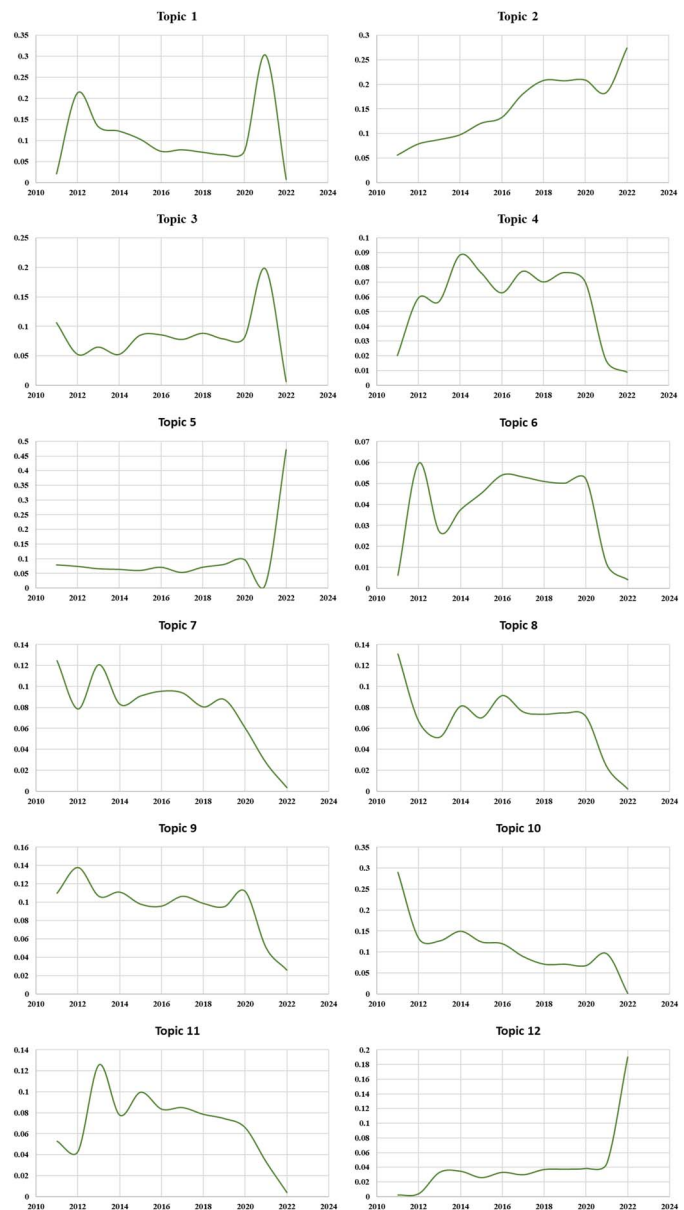

Fig. 7.   Digital solutions trends over time.

communications, and healthcare. By analyzing the changes in these topics over time, we aim to gain insights into the dynamics of these industries and the factors driving their growth, decline, or stability. This analysis will be useful for businesses seeking to identify opportunities for growth and make informed decisions about investments and strategies. To visualize the evolution of the proportion of each topic over time, we generated a smoothed curve depicting the proportion of each topic relative to the year of publication of the company's annual report, as illustrated in Fig. 7.

In analyzing the trends of the twelve topics over the years 2011 to 2022, it appears that there are three main groups: those with an overall upward trend (topics 2, 5, and 12), those with an overall downward trend (topics 7, 8, 9, 10, and 11), and those with a fluctuating trend (topics 1, 3, 4, and 6). However, there are some nuances within each group that are worth exploring.

1) Upward Trend: Emergence of New Technologies and Markets Driving Digitalization: The first group of topics, consisting of topics 2, 5, and 12, have exhibited an overall upward trend in terms of attention and discussion over the past few years. This trend could be attributed to the increasing importance of digital technologies and solutions in the business and healthcare sectors, as well as the growing demand for digital consulting and transformation services. Additionally, the global nature of these topics suggests the impact of globalization and the interconnectedness of different industries and markets.

However, it is interesting to note that topics 2 and 5 experienced a temporary dip in attention around 2021–2022 before resuming their upward trajectory. This could be due to several factors, including increased competition and saturation in the digital consulting and transformation space, as well as the impact of global events such as the COVID-19 pandemic. The pandemic has highlighted the need for digital solutions and technologies in healthcare and business, which could explain the subsequent increase in attention toward topic 12 in 2021.Overall, these trends suggest that the digital revolution is continuing to gain momentum, with businesses and healthcare providers increasingly adopting digital technologies and solutions. The impact of global events, such as the COVID-19 pandemic, further underscores the need for digital solutions and technologies, and the importance of keeping up with the latest trends in this rapidly evolving field.

2) Declining Trend: Economic Challenges and Competition Leading to a Shift Away From Digital Solutions: The second group of topics, consisting of topics 7–1, shows a declining trend in terms of attention and discussion. This trend may be indicative of a shift in focus from these specific digital solutions toward other emerging technologies and industries. One possible explanation for this shift is the increasing saturation and competitiveness in the digital marketplace, which is driving companies to explore alternative areas of growth and innovation. Additionally, global economic challenges and political instability may be contributing to a decrease in investment and interest in digital solutions, particularly in industries such as retail and entertainment that are heavily reliant on consumer spending. Despite these challenges, it is important for businesses to continue to prioritize digitalization and seek out new opportunities for growth and transformation in order to remain competitive in an increasingly digital world.

Topic 7, which focuses on digital commerce sales reporting, may be experiencing a decline in attention due to a shift toward more sophisticated data analytics and measurement tools that provide more comprehensive insights into consumer behavior and purchasing patterns. Similarly, topic 8, which centers around omnichannel media planning, may be facing competition from emerging technologies and platforms that provide more targeted and personalized marketing solutions. Topic 9, digital competition, may be seeing a decline in attention due to a growing focus on collaboration and partnerships between companies in the digital space, as well as increasing emphasis on customer-centric approaches to innovation. Topic 10, digital entertainment content distribution, may be experiencing a decline in interest due to changes in consumer behavior and preferences, such as

the rise of streaming services and the decline of traditional cable and broadcast television. Finally, topic 11, wireless communications technology and patent innovation, may be facing challenges due to the increasing complexity and regulation of the telecommunications industry, as well as competition from emerging technologies such as 5G and the Internet of Things.

3) Fluctuating Trend: Market Volatility and Changing Consumer Behaviors Impacting Digital Strategies: Topics 1, 3, 4, and 6 display fluctuating trends over the period from 2011 to 2022. Topics 1 and 3 show an upward trend followed by a decline in 2021–2022, while topics 4 and 6 exhibit a clear downward trend in the last two years. These trends can be attributed to a combination of factors related to the U.S. business environment, global economic conditions, and digitalization trends.

Topic 1, which focuses on digital solutions for surveying and data analysis across various sectors, may have experienced an initial increase in interest due to the growing importance of data-driven decision-making in business. However, the subsequent decline in attention may be due to the increasing availability of more advanced and specialized data analytics tools and solutions, which have emerged in recent years. Additionally, the impact of the COVID-19 pandemic on various industries may have reduced demand for these types of digital solutions. Topic 3, which concerns omnichannel retail strategies for seamless customer experience, saw a similar trend. The initial increase in attention can be attributed to the growing importance of online and mobile commerce and the need for retailers to provide a seamless customer experience across all channels. However, the subsequent decline may be due to the increasing saturation of the e-commerce market and growing competition from new entrants, as well as changes in consumer behavior and preferences. Topic 4, which pertains to corporate financing through the issuance of senior unsecured notes, may have experienced a decline in attention due to changes in the regulatory environment and the impact of global economic conditions, particularly the low-interest-rate environment. Additionally, the growing popularity of alternative financing options such as crowdfunding and venture capital may have reduced demand for traditional debt financing solutions. Topic 6, which deals with digital payments in the financial industry, may have experienced a decline in interest due to the increasing maturity and consolidation of the digital payments market, as well as the growing popularity of cryptocurrencies and blockchain-based payment solutions. The impact of regulatory changes and growing competition from nontraditional financial services providers may also have contributed to the decline.

Indeed, the dynamic shifts in digital topics over time highlight the importance of not only focusing on technology but also on management and strategic decision-making. As companies continue to undergo digital transformation, it is crucial to not only adopt new technologies but also to effectively manage the organizational changes that come with it. This includes understanding the changing needs and trends in the industry and adapting to them in a timely manner. Therefore, companies should focus on developing effective digital management strategies and frameworks that enable them to leverage digital technologies to meet their goals and objectives while remaining adaptable to changing circumstances.
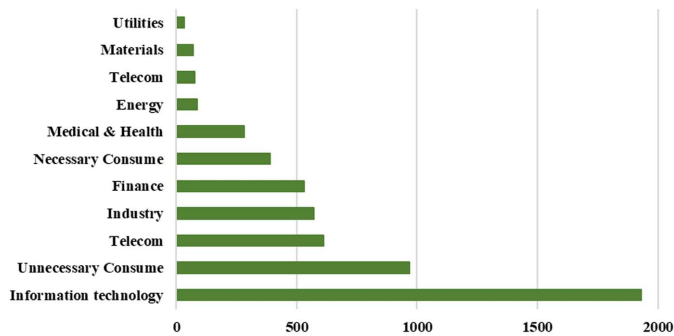
Fig. 8. Number of sections after filtering by industry.

Fig. 9. Information technology: topic proportions.

Fig. 10. Unnecessary consume: topic proportions.

C. Analysis of the Industry

In this section, we will explore the differences in digital transformation strategies and actions taken by companies across 11 major industries classified by the Global Industry Classification Standard (GICS). Industry classification research is essential for understanding the unique challenges and opportunities facing companies within each sector, as well as the specific digital solutions and technologies that are most relevant and effective in driving growth and innovation. By analyzing the digitalization trends and patterns within each industry, we can gain valuable insights into the current state of digital transformation and the future direction of the digital economy.

In order to explore the varying degrees of emphasis on digital transformation across different industries, we have filtered and identified the number of relevant paragraphs in the annual reports of each industry. The assumption is that the more paragraphs related to digital transformation and innovation, the greater the emphasis on digitalization in that industry. Based on this filtering process, we present in Fig. 8 the remaining number of paragraphs for each of the eleven GICS industries. This figure provides an initial overview of the differences in attention and prioritization of digital transformation across industries and sets the stage for further analysis.

Fig. 8 shows the number of remaining paragraphs after filtering for each industry, with information technology having the highest number of 1932 and utilities having the lowest number of 32. We hypothesize that industries with a higher number of remaining paragraphs are more likely to frequently mention digital innovation in their annual reports. This can be due to several factors, such as the industry's competitiveness and the need to stay up-to-date with technological advancements, regulatory pressures, and customer demand for digital solutions. Furthermore, the information technology and unnecessary consume industries, which have the highest number of remaining paragraphs, are typically more digitally focused and innovative than other industries, which could explain their higher frequency of mentioning digital innovation. Conversely, utilities and materials, which have the lowest number of remaining paragraphs, are traditionally less digitized and may have less pressure to invest in digital transformation.

In this section, we will conduct a topic proportion analysis on these two industries to gain a deeper understanding of their digital transfor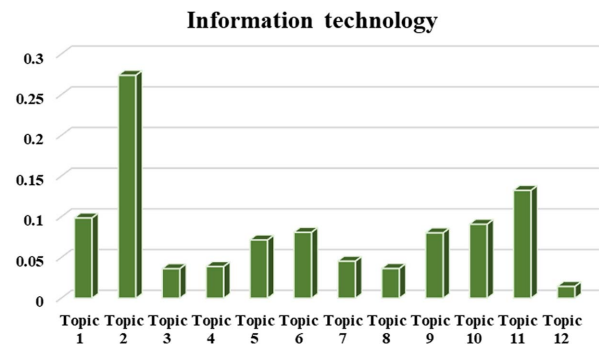mation strategies. While there are 11 industries in total, we choose to focus on the top two industries due to the limited space and the potential for more significant insights. By examining the topic proportions of these two industries, we can gain a better understanding of how digital transformation and innovation are perceived and executed in these sectors, represented in Figs. 9 and 10, respectively.
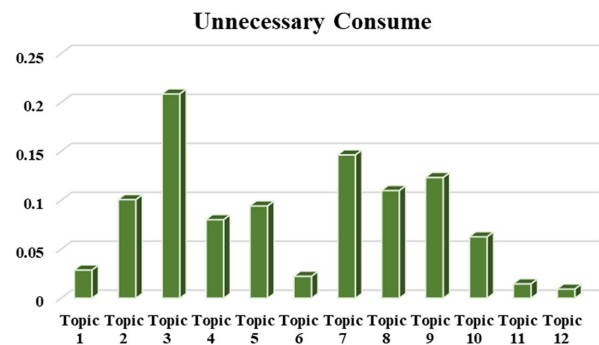
The information technology industry has a particularly high-topic proportion in topic 2, which represents "cognizant consulting and technology solutions for digital transformation." This high proportion is not surprising given the nature of the industry, which is driven by technological innovation and digital transformation. The high proportion suggests that companies in this industry are placing a strong emphasis on digital transformation and seeking consulting and technology solutions to facilitate this process. Companies that prioritize investing in digital solutions will likely have a competitive advantage in the future, as digital transformation is becoming increasingly critical for businesses to remain relevant and competitive.

On the other hand, the unnecessary consume industry has a particularly high-topic proportion in topic 3, which represents "omnichannel retail strategy for seamless customer experience." This high proportion is consistent with the industry's emphasis on customer experience and convenience. Companies in this industry are likely focused on improving their customer experience through the integration of online and offline sales channels, as well as providing customers with a seamless shopping experience across different platforms. To maintain a competitive edge in this industry, companies should continue to prioritize investing in and refining their omnichannel retail strategy, as customers are increasingly demanding a seamless shopping experience.

In conclusion, these topic proportion analyses suggest that companies in the information technology industry are prioritizing digital transformation through consulting and technology solutions, while companies in the unnecessary consume industry are prioritizing customer experience through omnichannel retail strategies. Both strategies are critical for companies in these industries to remain competitive in the rapidly evolving business landscape. Companies should continue to invest in and refine these strategies while also exploring other innovative approaches to further improve their business operations.

V. CONCLUSION

In this article, we have introduced a valuable novel metric for assessing innovation in enterprise digital transformation, utilizing textual analysis of annual reports from publicly listed companies in the United States. Our text-based metrics offer a comprehensive depiction of the digital innovation strategies of these firms. This research has contributed significantly to the theoretical significance of the fields of enterprise management and sustainable development.

First, this study answers several key questions about the digital innovation strategy of enterprises. Our research results indicate that the prevalence of digital transformation, competition, and risk management in business operations is a crucial issue for companies, and external support is necessary to address these challenges. Through network analysis and topic modeling, we identified 12 topics related to digital transformation in annual reports of U.S.-listed companies. These topics can be grouped into three main categories: digital technology innovation, customer-oriented digital strategy, and digital transformation in traditional business operations. By leveraging these insights, organizations can adopt digital tools and strategies to enhance their operations, gain a competitive advantage, and manage digital risks effectively.

Second, it can be observed that from 2011 to 2022, there has been a significant increase in the number of annual reports from U.S.-listed companies that focused on digital transformation and innovation strategies. This can be attributed to the growing recognition of the importance of digital transformation in maintaining a competitive edge and adapting to the ever-changing market demands. Notably, the focus of enterprise digital innovation strategy has shifted over the 12-year period from a technology-centric approach to a more innovation management-oriented approach. These trends indicate the ever-increasing momentum of the digital revolution, and the pressing need for enterprises to prioritize digitization to sustain their competitiveness in an increasingly digital world.

Third, the analysis of topic proportions suggests that companies in different industries have varying priorities when it comes to digital innovation and transformation. However, it is clear that digital transformation is becoming increasingly critical for businesses to remain relevant and competitive, regardless of industry. Therefore, companies should continue to invest in digital solutions and refine their strategies to stay ahead of the curve. Additionally, regulatory pressures and customer demand for digital solutions are likely to continue driving the adoption of digital innovation across all industries in the future.

In summary, our research introduces a new metric for measuring innovation in digital transformation of enterprises, making a significant contribution to the field of business management and sustainability. Its impact extends beyond academia, providing valuable insights for organizations, governments, and society.

For organizations, our research provides important insights into digital innovation strategies, highlighting the importance of digital transformation, competition, and risk management. By identifying key digital transformation topics in U.S. public company annual reports, we provide guidance to organizations to enhance their operations, gain a competitive advantage, and effectively manage digital risks to make informed decisions in an increasingly digital environment.

On a broader scale, our findings highlight the growing importance of digital transformation in staying competitive and adapting to market dynamics. The shift toward an innovation management orientation reflects the dynamism of the digital revolution. Policymakers can use these insights to develop strategies that support digital innovation, drive economic growth, and enhance economic resilience. Our research highlights the pervasive relevance of digital transformation across industries, highlighting the need for continued investment in digital solutions and strategic improvements to be of inestimable value to organizations, policymakers, and society embracing the digital revolution.

While our research has provided insightful findings, there remain certain limitations that can be addressed in future research. First, the veracity of the content within the corporate annual reports analyzed may be subject to questioning. Some companies may overemphasize or downplay their digital innovation strategies [49], which may impact the validity of the conclusions drawn. Future studies may mitigate this issue by gathering data from alternative sources. Second, certain data points were not available for analysis. The majority of digital innovation strategies retrieved for this study were sourced from company disclosures, which may be influenced by various factors that could impact the findings [50].

Finally, while our research focus is on corporate digital transformation strategies, we relied solely on textual data for analysis. Future studies may incorporate other data sources to provide a more comprehensive assessment of the impact of digital innovation strategies on corporate performance. Our study highlights the potential for exploring textual information as a valuable avenue for future research.

REFERENCES

[1] G. Vial, "Understanding digital transformation: A review and a research agenda," *J. Strategic Inf. Syst.*, vol. 28, no. 2, pp. 118–144, 2019.

[2] G. Mulgan, S. Tucker, R. Ali, and B. Sanders, "Social innovation: What it is, why it matters and how it can be accelerated," in *Skoll Centre for Social Entrepreneurship*, London, U.K.: Young Foundation, 2007.

[3] C. Matt, T. Hess, and A. Benlian, "Digital transformation strategies," *Bus. Inf. Syst. Eng.*, vol. 57, pp. 339–343, Aug. 2015.

[4] P. Stanton and J. Stanton, "Corporate annual reports: research perspectives used," *Account. Audit. Account. J.*, vol. 15, no. 4, pp. 478–500, 2002.

[5] A. Di Vaio et al., "Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review," *J. Bus. Res.*, vol. 121, pp. 283–314, Dec. 2020.

[6] F. Lüdeke-Freund, "Sustainable entrepreneurship, innovation, and business models: Integrative framework and propositions for future research," *Bus. Strategy Environ.*, vol. 29, no. 2, pp. 665–681, 2020.

[7] H. Schaffers et al., *Smart Cities and the Future Internet: Towards Cooperation Frameworks for Open Innovation*. Berlin, Germany: Springer-Verlag, 2011.

[8] I. Guandalini, "Sustainability through digital transformation: A systematic literature review for research guidance," *J. Bus. Res.*, vol. 148, pp. 456–471, Sep. 2022.

[9] R. Busulwa, M. Pickering, and I. Mao, "Digital transformation and hospitality management competencies: Toward an integrative framework," *Int. J. Hospitality Manage.*, vol. 102, 2022, Art. no. 103132.

[10] N. M. Bocken et al., "A literature and practice review to develop sustainable business model archetypes," *J. Cleaner Prod.*, vol. 65, pp. 42–56, Feb. 2014.

[11] F. Piller, A. Vossen, and C. Ihl, "From social media to social product development: The impact of social media on co-creation of innovation," Die Unternehmung-Swiss, *J. Bus. Res. Practice*, vol. 66, no. 1, pp. 7–27, 2012.

[12] A. Miklosik and N. Evans, "Impact of big data and machine learning on digital transformation in marketing: A literature review," *IEEE Access*, vol. 8, pp. 101284–101292, 2020.

[13] F. Westley and N. Antadze, "Making a difference: Strategies for scaling social innovation for greater impact," *Innov. J.*, vol. 15, no. 2, pp. 289–310, 2010.

[14] J. A. Phills, K. Deiglmeier, and D. T. Miller, "Rediscovering social innovation," *Stanford Social Innov. Rev.*, vol. 6, no. 4, pp. 34–43, 2008.

[15] N. Cornelius et al., "Corporate social responsibility and the social enterprise," *J. Bus. Ethics*, vol. 81, pp. 355–370, Jul. 2008.

[16] S. Romero, S. Ruiz, and B. Fernández-Feijóo, "Assurance statement for sustainability reports: The case of Spain," in *Proc. Northeast Bus. Econ. Assoc.*, pp. 105–112, 2010.

[17] Y. Xu et al., "Does digital transformation foster corporate social responsibility? Evidence from Chinese mining industry," *J. Environ. Manage.*, vol. 344, 2023, Art. no. 118646.

[18] J. Bughin et al., *Artificial Intelligence the Next Digital Frontier*, San Francisco, CA, USA: McKinsey Global Institute, 2017.

[19] P. M. Bican and A. Brem, "Digital business model, digital transformation, digital entrepreneurship: Is there a sustainable 'digital'?" *Sustainability*, vol. 12, no. 13, 2020, Art. no. 5239.

[20] T. Hautala-Kankaanpää, "The impact of digitalization on firm performance: examining the role of digital culture and the effect of supply chain capability," *Bus. Process Manage. J.*, vol. 28, no. 8, pp. 90–109, 2022.

[21] D. Proksch et al., "The influence of a digital strategy on the digitalization of new ventures: The mediating effect of digital capabilities and a digital culture," *J. Small Bus. Manage.*, pp. 1–29, Mar. 2021.

[22] H. Fromm, F. Habryn, and G. Satzger, "Service analytics: Leveraging data across enterprise boundaries for competitive advantage," in *Globalization of Professional Services: Innovative Strategies, Successful Processes, Inspired Talent Management, and First-Hand Experiences*, pp. 139–149, Jan. 2012.

[23] M. Fitzgerald et al., "Embracing digital technology: A new strategic imperative," *MIT Sloan Manage. Rev.*, vol. 55, no. 2, pp. 1–12, 2014.

[24] P. J. Williamson and A. De Meyer, "Ecosystem advantage: How to successfully harness the power of partners," *California Manage. Rev.*, vol. 55, no. 1, pp. 24–46, 2012.

[25] J. B. A. P. Filho, T. B. Lacerda, and S. R. de L. Meira, "Digital [business|transformation] strategy: A systematic review: Could companies benefit from a consolidated road to transformation?" in *Proc. 17th Iberian Conf. Inf. Syst. Technol. (CISTI)*, 2022, pp. 1–7.

[26] E. Avdeeva, T. Averina, and S. Barkalov, "Innovative models and strategies for digital transformation of companies for sustainable development," in *Proc. 3rd Int. Conf. Control Syst., Math. Model., Autom. Energy Efficiency (SUMMA)*, 2021, pp. 340–343.

[27] S. Yucel, "Modeling digital transformation strategy," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, 2018, pp. 221–226.

[28] Z. Lei and W. Zhaolei, "The realization path of digital transformation of China's economic development based on SWOT analysis," in *Proc. Int. Conf. Public Manage. Intell. Soc. (PMIS)*, 2021, pp. 115–118.

[29] N. Behare, S. Waghulkar, and S. A. Shah, "A theoretical perspective on customer experience (CX) in digital business strategy," in *Proc. Int. Conf. Res. Intell. Comput. Eng. (RICE)*, 2018, pp. 1–7.

[30] I. Vayansky and S. A. Kumar, "A review of topic modeling methods," *Inf. Syst.*, vol. 94, 2020, Art. no. 101582.

[31] R. Niveditha et al., "Develop CSR themes using text-mining and topic modelling techniques," in *Proc. IEEE Int. Conf. Cloud Comput. Emerg. Markets (CCEM)*, 2020, pp. 67–71.

[32] M. Jagannathan, D. Roy, and V. S. K. Delhi, "Application of NLP-based topic modeling to analyse unstructured text data in annual reports of construction contracting companies," *CSI Trans. ICT*, vol. 10, no. 2, pp. 97–106, 2022.

[33] A. H. Huang et al., "Analyst information discovery and interpretation roles: A topic modeling approach," *Manage. Sci.*, vol. 64, no. 6, pp. 2833–2855, 2018.

[34] G. Bellstam, S. Bhagat, and J. A. Cookson, "A text-based analysis of corporate innovation," *Manage. Sci.*, vol. 67, no. 7, pp. 4004–4031, 2021.

[35] K. Li et al., "Measuring corporate culture using machine learning," *Rev. Financial Stud.*, vol. 34, no. 7, pp. 3265–3315, 2021.

[36] N. Pröllochs and S. Feuerriegel, "Business analytics for strategic management: Identifying and assessing corporate challenges via topic modeling," *Inf. Manage.*, vol. 57, no. 1, 2020, Art. no. 103070.

[37] S. Bai et al., "What makes a theme park experience less enjoyable? Evidence from online customer reviews of Disneyland China," *Front. Psychol.*, vol. 14, pp. 103–113, Jun. 2023.

[38] S. Bai et al., "Exploring thematic influences on theme park visitors' satisfaction: An empirical study on Disneyland China," *J. Consum. Behav.*, vol. 23, no. 1, pp. 90–106, 2024.

[39] B. Chae and D. L. Olson, "Discovering latent topics of digital technologies from venture activities using structural topic modeling," *IEEE Trans. Comput. Social Syst.*, vol. 8, no. 6, pp. 1438–1449, Dec. 2021.

[40] Z. S. Harris, "Distributional structure," *Word*, vol. 10, nos. 2–3, pp. 146–162, 1954.

[41] T. Mikolov et al., "Efficient estimation of word representations in vector space," 2013, *arXiv:1301.3781.*

[42] M. E. Roberts et al., "Structural topic models for open-ended survey responses," *Amer. J. Political Sci.*, vol. 58, no. 4, pp. 1064–1082, 2014.

[43] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Jan. 2003.

[44] M. E. Roberts et al., "The structural topic model and applied social science," in *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation*, vol. 4, no. 1, Cambridge, MA, USA: Harvard University, 2013.

[45] J. Büschken and G. M. Allenby, "Sentence-based text analysis for customer reviews," *Marketing Sci.*, vol. 35, no. 6, pp. 953–975, 2016.

[46] M. E. Roberts, B. M. Stewart, and D. Tingley, "STM: An R package for structural topic models," *J. Statist. Softw.*, vol. 91, pp. 1–40, Oct. 2019.

[47] M. E. Roberts, B. M. Stewart, and E. M. Airoldi, "A model of text for experimentation in the social sciences," *J. Am. Stat. Assoc.*, vol. 111, no. 515, pp. 988–1003, 2016.

[48] K. Li et al., "Stable matching-based selection in evolutionary multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 18, no. 6, pp. 909–923, Dec. 2013.

[49] E.-H. Kim and T. P. Lyon, "Greenwash vs. brownwash: Exaggeration and undue modesty in corporate sustainability disclosure," *Org. Sci.*, vol. 26, no. 3, pp. 705–723, 2015.

[50] S. Cowan and D. Gadenne*, "Australian corporate environmental reporting: A comparative analysis of disclosure practices across voluntary and mandatory disclosure systems," *J. Account. Organ. Change*, vol. 1, no. 2, pp. 165–179, 2005.

Shizhen Bai received the Ph.D. degree in system engineering from Southeast University, China, in 2007.

He is a Professor and a Doctoral Supervisor with the School of Management, Harbin University of Commerce, Harbin, China. His main research interests include service management, big data analysis, etc. His research results have been published in *Electronic Commerce Research and Applications, Journal of Intelligent Manufacturing, Frontiers in Psychology, Frontiers of Engineering Management*, among others. He has published many monographs in supply chain management and service management.

Yongbo Tan is working toward the master's degree in management science and engineering with the School of Management, Harbin University of Commerce, Harbin, China.

He possesses a robust skill set in machine learning, text analysis, and data processing. His research interests include enterprise innovation management and supply chain management.
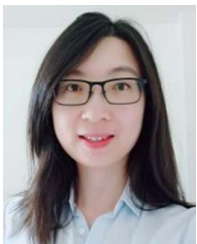
Chunjia Han received his Ph.D. degree in management.

He is a Senior Lecturer (Associate Professor) in business innovation with Birkbeck, University of London, London, U.K. He is a Visiting Professor and the Ph.D. Supervisor at a number of universities (e.g., Beijing Jiaotong University, Beijing, China and Sichuan University, Chengdu, China). By focusing on the application, development, and impact of digital technologies (e.g., big data analytics, AI, and social media) in influencing practices in business and economy, his research contributes to our knowledge on innovation management, tourism management, digital marketing, and digital economy. He is the Author or a Co-Author of over 30 publications. His research has appeared in journals, including IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT, *International Journal of Product Economics*, *Technological Forecasting and Social Change*, *Annals of Tourism Research*, *Journal of Business Research*, *Asia Pacific Journal of Management*, among others.

Dr. Han has been serving as a Guest Editor of special issue for 14 academic journals, including *Journal of Global Information Management*, *Journal of Advanced Transportation*, *International Journal of Engineering Business Management*, and *Computer Modeling in Engineering & Sciences.* He has co-organised conferences and served as a Committee Member for conferences/workshops, including IEEE International Conference on Bioinformatics and Biomedicine, IEEE International Workshop on Machine Learning for Trust, Security and Privacy in Computing and Communications, IEEE International Conference on Consumer Electronics, and The Web Conference. He currently serves as the Deputy Director of the International Creative Management Committee.

Mu Yang received the Ph.D. degree in computer science from the University of Southampton, in 2014.

She is a Senior Lecturer (Associate Professor) in business analytics with Birkbeck, University of London, London, U.K. Her research interests include area of digital innovation, marketing, social media research, and user privacy. Her research typically uses big data analytics and user behavior modeling approaches.

Brij B. Gupta (Senior Member, IEEE) received his Ph.D. degree in computer science and engineering.

He is the Director of International Center for AI and Cyber Security Research and Innovations, Taichung, Taiwan, and a Distinguished Professor with the Department of Computer Science and Information Engineering (CSIE), Asia University, Taichung, Taiwan. In his professional experience of more than 18 years, he published over 500 papers in journals/conferences, including 35 books and 12 patents with over 25 000 citations. His research interests include information security, cyber–physical systems, cloud computing, blockchain technologies, intrusion detection, AI, social media, and networking.

Dr. Gupta has received numerous national and international awards, including Canadian Commonwealth Scholarship (in 2009), Faculty Research Fellowship Award (in 2017), MeitY, GoI, IEEE GCCE outstanding, and WIE paper awards and the Best Faculty Award (in 2018 and 2019, respectively), NIT KKR. He was selected for Clarivate Web of Science Highly Cited Researchers in computer science (as one of the top 0.1% researchers in the world) consecutively in 2022 and 2023. He was selected as one of the world's top 2% scientists in Stanford University's ranking consecutively in 2020, 2021, 2022, and 2023. He is a Visiting/Adjunct Professor with several universities worldwide. He is a Distinguished Lecturer in IEEE CTSoc, in 2021. He is a Member-in-Large, Board of Governors, IEEE Consumer Technology Society (2022–2024). He is the Editor-in-Chief of International Journal on Semantic Web and Information Systems (IJSWIS), International Journal of Software Science and Computational Intelligence (IJSSCI), STE, and International Journal of Cloud Applications and Computing (IJCAC). Moreover, he is the Lead-Editor of a book series with CRC and IET Press. He served as a TPC Member in more than 150 international conferences. He is an Associate/Guest Editor of various journals and transactions.

Varsha Arya received the master's degree in department of business administration from Rajasthan University, Jaipur, India, in 2015. Currently, she is working toward the Ph.D. degree in department of business administration with Asia University, Taichung, Taiwan, China.

She has been working as a Researcher for the last seven years in different organizations. She published more than 25 papers in top journals and conferences. Her research interests include business administration, technology management, cyber–physical systems, cloud computing, healthcare, and networking.

Neeraj Kumar (Senior Member, IEEE) received his Ph.D. degree in computer science and engineering.

He is currently a Full Professor with the Department of Computer Science and Engineering, Thapar Institute of Engineering and Technology (Deemed to be University), Patiala, India. He has published more than 400 technical research papers in top-cited journals and conferences, which are cited more than 29 410 times from well-known researchers across the globe with a current H-index of 93. He has guided many research scholars toward the Ph.D. and M.E./M.Tech. degrees. His research was supported by various competitive agencies across the globe. His research interests include green computing and network management, IoT, big data analytics, deep learning, and cybersecurity. He has edited/authored ten books with international/national publishers like IET, Springer, Elsevier, and CRC, such as *Security and Privacy of Electronic Healthcare Records: Concepts, Paradigms and Solutions*, *Machine Learning in Cognitive IoT* (CRC Press), *Blockchain, Big Data and Machine Learning* (CRC Press), *Blockchain Technologies Across Industrial Vertical* (Elsevier), *Multimedia Big Data Computing for IoT Applications: Concepts, Paradigms and Solutions* by the Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S, 2019), and *Probabilistic Data Structures for Blockchain-Based Internet of Things Applications* (CRC Press). One of the edited book titled *Multimedia Big Data Computing for IoT Applications: Concepts, Paradigms and Solutions* (Springer, 2019) has 3.5 million downloads until June 2020. It attracts the attention of researchers across the globe. He is a Highly Cited Researcher from WoS, from 2019 to 2021.