A Comprehensive Bibliometric Analysis of Big Data in Entrepreneurship Research

Anran Xiao[1], Yong Qin[1], Zeshui Xu[1]*, Marinko Skare[2,3]

[1]*Business School, Sichuan University*
Chengdu 610064, China
E-mail: xiaoanran_ubby@163.com; yongqin_ahsc@163.com; xuzeshui@263.net
*Corresponding author

[2]*Juraj Dobrila Univ Pula, Fac Econ & Tourism Dr Mijo Mirkovic*
Preradoviceva 1-1, Pula 52100, Croatia
[3]*University of Economics and Human Sciences Warsaw, Poland*
E-mail: mskare@unipu.hr

Big data technology has been widely used in entrepreneurial research in recent years. To explore the development trend and fundamental characteristics of big data in entrepreneurship (BDIE) research, we conduct a comprehensive bibliometric analysis of BDIE based on 541 publications between 1993 and 2020 from Web of Science. On the one hand, this paper focuses on some essential characteristics of the BDIE publications, such as categories, citation, H-index, and the most cited publications. On the other hand, visual scientific maps are presented by bibliometric tools, i.e., Bibliometrix, VOS viewer, and CorTexT Manager, showing relationships between publications and knowledge structure in BDIE research. Finally, hot topics in current studies, knowledge, and limitations are discussed, guiding scholars to explore new research directions. This paper provides a relatively broad perspective for applying BDIE research, which contributes to understanding the evolution of BDIE research and inspires scholars in related fields.

Keywords: *Big Data in Entrepreneurship; Bibliometric Analysis; CorTexT Manager; Science Mapping.*

Introduction

Big data analytics is a cutting-edge technology that is transforming the world we live in. By powering the development of information technologies and shaping the way people think and behave, it's impact and applications are undeniable and far-reaching. With the ability to provide crucial new insights into innovation development and sustainable entrepreneurship, big data analytics has become an essential tool for businesses looking to stay ahead of their competitors. As a result, many firms are investing heavily in big data to uncover innovative ways to differentiate themselves and gain a competitive edge in the market (Corte-Real *et al*., 2017). The consensus among experts is that big data is set to become a crucial driver of innovation, competition, and productivity in the near future, representing the next major frontier in technology (Manyika *et al*., 2011; MaryAnne *et al*., 2015). Big data, characterized by variety, velocity, and volume of data (Artz, 2013), can change the innovation landscape by increasing the fit between consumer preferences and product features (Chen *et al*., 2021; Lei *et al*., 2019; Pogrebna, 2015). The distinct characteristics of big data, including its volume, variety, and velocity, have a profound impact on a business's innovation performance, shaping customer perspectives, financial returns, and operational excellence. Therefore, the capacity to effectively harness big data is critical for businesses seeking to achieve success and gain a competitive advantage (Calic & Ghasemaghaei, 2020). By examining the development of big data research technology, businesses and government organizations can enhance their technology

management, streamline operations, and uncover fresh prospects for research and development. This analysis offers valuable insights into the growth of big data and its applications, enabling organizations to stay ahead of the curve and leverage new opportunities for innovation (Ciampi *et al*., 2020; Park, 2021). Big data is an indispensable catalyst for the growth of emerging technologies and is intimately connected to various technological innovations in data processing and storage. (Manogaran *et al*., 2016; Zhou *et al*., 2014). The rise of big data analytics has created opportunities for entrepreneurs to gain valuable insights and make data-driven decisions using hardware and software techniques to analyze large, complex datasets (Manogaran *et al*., 2016).

Entrepreneurship in a given country can have a targeted impact on specific areas, including economic growth, productivity, technology innovation, new product development, market disruption, foreign trade, optimal use of domestic resources, inter-firm collaboration, and equitable distribution of wealth within society (Akpor-Robaro *et al*., 2012). Entrepreneurship is seen as an endogenous stimulus for business development in many developed economies (Zeng, 2018). To some extent, innovation and initiative are particularly important for business operation in an entrepreneurial orientation (Lipych *et al*., 2021). On the other hand, with the rapid development of the big data internet system, the development and application of big data in entrepreneurship are of great interest to scholars. BDIE means big data in entrepreneurship, which covers the application and development of big data in innovation and entrepreneurship,

with the value and role of big data in entrepreneurship being worth exploring. Big data is widely used in the fields of finance (Pan, *et al.*, 2021), education (Hao, *et al.*, 2021; Xie, 2021) and management for entrepreneurship. Big data-supported technologies such as text analytics, machine learning techniques and digitization facilitate the reduction of entrepreneurial risk (Zhou & Wang, 2021), prediction of innovation performance (Zheng *et al.*, 2021), more effective entrepreneurial models (Ferrati & Muffatto, 2021), and the generation of entrepreneurial spirit and mindsets (Zurita *et al.*, 2021). Specifically, entrepreneurial ventures are dynamic processes that face interactive and diverse potential risks, and big data analytics can effectively enable integrated management of risk factors and reduce entrepreneurial risk (Zhou & Wang, 2021). For example, big data can be used for text analysis to predict the success of projects based on language preferences (Du *et al.*, 2021). Furthermore, machine learning techniques with data-driven characteristics are also applied to demonstrate and predict the performance of open innovation (Zheng *et al.*, 2021), which contribute to investment decisions and can provide new insights on models that are more effective in entrepreneurship (Ferrati & Muffatto, 2021). In addition, big data enabling (BDE) and human resource management (HRM) can inspire organizational dynamics and positively impact employee intrapreneurship. In turn, employee intrapreneurship synergy between BDE, business performance, and HRM (Wan & Liu, 2021). Digital transformation companies emphasize the use of big data to facilitate the generation of new mindsets (Zurita *et al.*, 2021). For example, owners of enterprises must ensure more robust investment in their social infrastructure through the development of digitalization and reputational capital (Urban & Mutendadzamera, 2021). Furthermore, big data is one of the innovative ways to help companies improve their supply chain agility (Kittichotsatsawat *et al.*, 2021; Wang & Ali, 2021).

To sum up, big data is applied in a wide range of fields and plays an important role in entrepreneurship. Some studies have used a systematic literature review approach to explore ways in which big data can foster innovation (Walter, 2021), and the potential of big data for entrepreneurship (Obschonka, 2017; Zulkefly *et al.*, 2021). In addition, the impact of digital tools such as online review platforms (Mariani & Nambisan, 2021) and customer relationship management systems (Guerola-Navarro *et al.*, 2022) on entrepreneurship has been widely discussed through literature review. However, the exciting research lacks a more objective and visual approach to exploring and discussing the development, collaborations, research themes, and future research directions of BDIE research. The number of research publications in BDIE is so sheer that there is no way to read each one in detail, as we do with a literature review. Furthermore, a literature review does not give a complete and objective picture of knowledge, as the purposeful selection of a non-random sample will bias the analysis (Linnenluecke *et al.*, 2020). On the one hand, a literature review is limited by time and content, and short-term studies cannot present trends over time. On the other hand, the selection and retelling of articles are somewhat subjective.

To tackle this problem, scholars require a thorough and systematic bibliographic analysis to comprehend the current state and limitations of BDIE research. a bibliometric analysis of BDIE research provides an informative overview of the available publications, which is a rapidly developing field that is making significant progress. Bibliometrics, quantitative analysis, and statistical evaluation (Meng & Chan, 2021) are evolving into a vital part of the decision-making process (Durana *et al.*, 2020) and are widely used in a variety of fields, such as management, entrepreneurship, and innovation (Baier-Fuentes *et al.*, 2019; Qin *et al.*, 2021). In this article, we explore new trends and provide visual and understandable scientific mapping while dealing with a large amount of data.

More precisely, the main contributions of this paper are as follows: (1) Performance analysis. It presents the development of the BDIE articles over the years and their future potential. At the same time, a review of highly cited publications provides insight into what is popular and highly valued in the BDIE research. (2) The analysis at the country/region and author level. It reveals highly productive authors and countries/regions and excavates authors with potential and energy contributing to the future development of the BDIE research. Meanwhile, international cooperation in BDIE research is also presented visually. (3) Keyword analysis. We delve into the content level of the BDIE research, showing the various clusters, the relevance of countries/regions and authors to the themes, and the thematic evolution. Therefore, the prospects and additional possibilities in BDIE research are explored in greater depth. (4) A comprehensive discussion. This paper discusses hot issues, enlightenment, and limitations based on the above analysis.

In the following sections, we illustrate the research framework of the methodology used in this study in Section 2. Furthermore, Section 3 is developed in three ways: performance analysis, countries/regions, and authors, and keyword analysis. In Section 4, we elaborate on the development prospects and limitations of the paper. Finally, we draw some conclusions to conclude the article.

Methodology

This study aims to provide a comprehensive and systematic bibliographic analysis of BDIE research. First, some essential characteristics of these publications are statistically analyzed to present the distribution over the years and explore future development trends in BDIE research. Subsequently, a collaborative analysis of countries/regions and authors is carried out through bibliometric tools, which vividly visualize the interrelationships in collaborative networks. Last but not least, keywords analysis suggests the intellectual structure and future research directions of the BDIE research, which contributes scholars to exploring new approaches and doing in-depth research. Figure 1 shows the research framework of the methodology used in this study.

Data Source and Search Protocol

In this paper, a bibliometric method is used to explore big data in entrepreneurship. Our data come from the Web of Sciences (WoS) core collection database owned by Thomson & Reuters Co. The WoS is one of the most credible indexes and databases (Farooq *et al.*, 2021) in the social sciences, especially in the field of business and

economics (Gonzalez-Torres *et al.*, 2020), because it not only has authoritative and high-quality publications but also provides reliable, detailed information of data sets for analysis (Wang *et al.*, 2021). To maintain data integrity, the search was carried out on October 13, 2021, and we chose the timespan from 1900 to 2020 for a more comprehensive data collection. After eliminating invalid data from the original dataset, a total of 541 publications from 1993 to

2020 are collected, which will be applied to analyze and explore new development patterns of BDIE research. The specific details of the search protocol are summarized as follows:

Topics = Big data and Entrepreneur*
Timespan = from 1900 to 2020
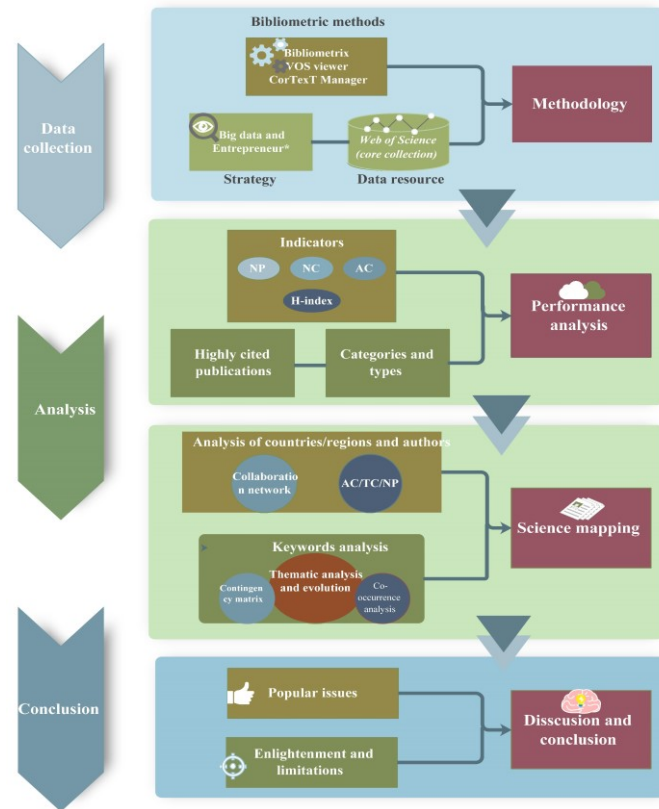**Database** = WoS Core Collection

Figure 1. The Research Framework of the Methodology used in this Study

Bibliometric Methods

Bibliometric analysis is conducted along two primary dimensions: performance analysis and science mapping, which analyzes the research performances of scholars, national/regional collaborations, the importance of themes, journals, countries, and so on (Gao *et al.*, 2021; Moral-Munoz *et al.*, 2020). On the one hand, performance analysis focuses on the basic information of the publications to evaluate the current development state and performance by observing the growth trend of numbers and citations, categories, and highly cited literature (Wang *et al.* 2020; Yu *et al.*, 2017). On the other hand, science mapping presents visual knowledge bases through cocitation, cooccurrence, and keyword analysis to show relationships among publications and explore the structure and evolution of themes in the field (Guleria & Kaur, 2021; Wang *et al.*, 2021).To conduct a more comprehensive and practical bibliometric analysis, we present visual and understandable performance analysis and science mapping by combining the advantages of three bibliometric tools, i.e., Bibliometrix (Aria & Cucurullo, 2017), VOS viewer (Vaneck & Waltman, 2010), and CorTexT Manager (http://manager.cortext.net/).

Results

Performance Analysis

(1) Annual indicators of publications

According to the data shown in WoS and the analysis results given by Bibliometrix, it is accessible to conduct a performance analysis, which contributes to learning the status and emerging trends of the BDIE research. BDIE research has been going on for 28 years, and essential information is shown in Table 1. It includes 541 independent studies on defined topics, with 1,365 authors applying 1,793 different keywords to categorize their research. The status of BDIE research is comprehensively analyzed from the perspective of the number of publications and citations, which contributes to observing the development trend. More publications are presented to explore which areas are of greater interest to the BDIE literature. In contrast, highly cited publications are studied to determine what topics are popular and relevant to scholars. The following three sections will elaborate on the parts mentioned above.

Table 1
Essential Information about BDIE Research

Description	Results
Timespan	1993–2020
Sources (Journals, Books, etc.)	416
Documents	541
Average years from publication	4
Average citations per document	10.51
Average citations per year per document (AC)	2.226
References	23,407
Keywords plus	973
Author keywords	1,793
Authors	1,365

In the following, we will analyze the annual indicators of publications. Figure 2 (a) presents the number of publications (NP) per year from 1993 to 2020. The entrepreneurship literature on big data has recently gained increasing attention from scholars. Based on the trend of NP in the column graph, we can divide the time into three periods to more clearly observe the changing trends and explore the development characteristics of the BDIE research.

I: 1993–2004 (start-up period). This phase is the beginning of the BDIE research. As a result, the number of publications is less than four per year and shows no sign of growth.

II: 2005–2012 (initial development period). At this period, BDIE research began to show signs of growth, with details finding that the number of articles posted at this stage was double that of the previous one, growing to around seven articles.

III: 2013–2020 (rapid development period). The number of articles exploded in 2013 and has continued to increase rapidly in the following years. It should be noted that the number of articles increased to more than 50 in 2016 and dramatically again to 122 in 2020.

For one thing, by observing the changing trends of NP over these three periods, it can be seen that big data research promotes the application of relevant technologies in innovation, and the number of BDIE research is expected to continue to grow. For example, the number of citations (NC) per year of BDIE research is presented in Figure 2 (b), and the NC that is less than 200 slowly increases in Period II. However, NC has been overgrowing in Period III and reached a peak of 1,340 in 2020, followed by 955 in 2019, which shows that BDIE research is becoming increasingly influential and attracting the attention of more scholars.
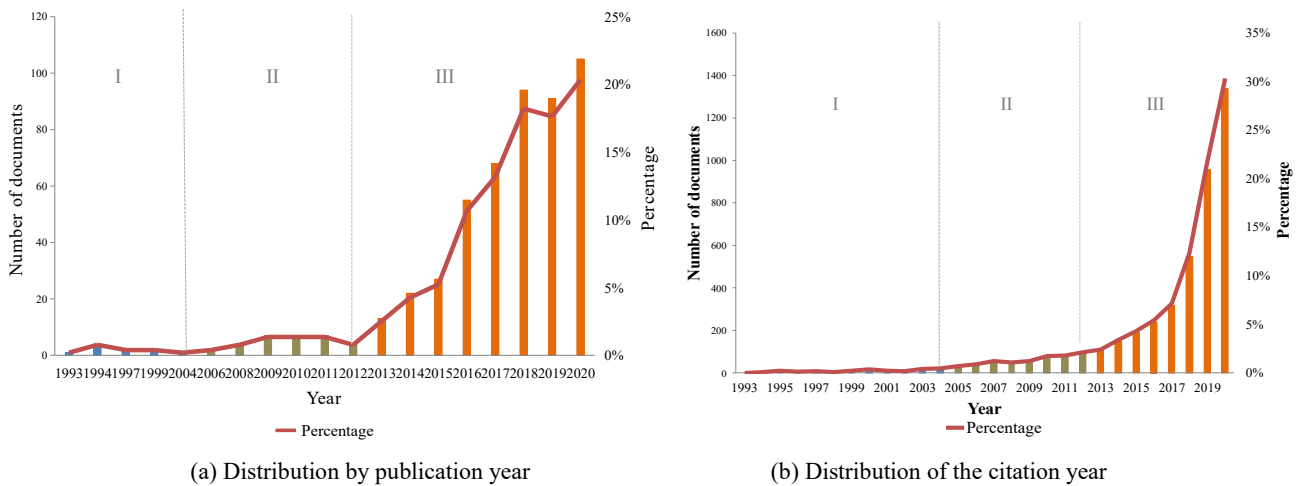

(a) Distribution by publication year　　　(b) Distribution of the citation year

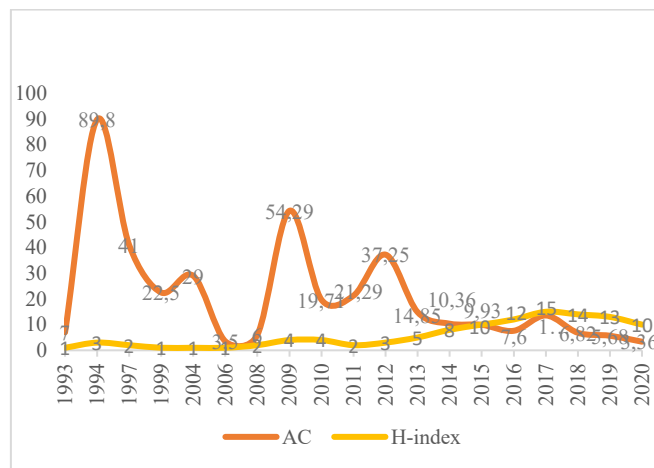**Figure 2.** The Publication Year and the Citation Year Distribution

Figure 3. The AC and H-Index of Publications

Furthermore, the h-index shown in Figure 3 gives a more accurate picture of academic achievement. The higher the h-index of the year, the more significant the impact of BDIE research. The average number of citations per year (AC) fluctuates considerably in Periods I and II and then plateaus as it steps into Period III. At first glance, it is clear that the AC is slowly decreasing in Period III, caused by a significant increase in the number of publications. However, AC fluctuations decrease with time, indicating that BDIE research is stabilizing and gaining more attention.

(2) Categories of Publications

Figure 4 shows the top 10 topics of the BDIE research in WoS. "Business" is the leading research direction among them, accounting for 18 %. Furthermore, management ranks second, accounting for 17 %. However, research directions, such as Computer Science Artificial Intelligence and Operations Research Management Science, have fewer publications.
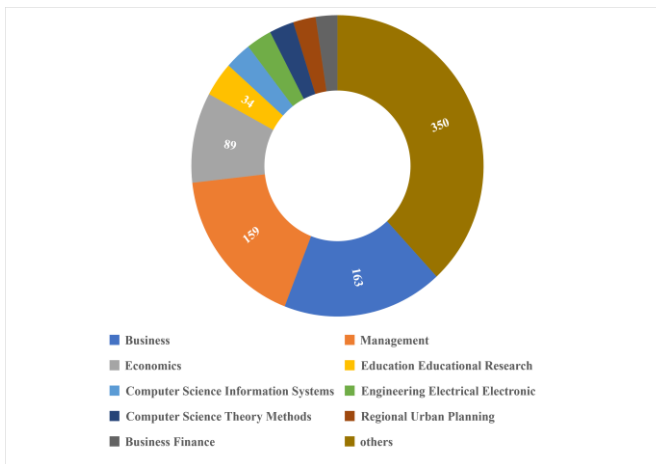
b) Distribution of categories

Figure 4. The Categories of Documents of the BDIE in WoS (1993–2020)

(3) Highly cited Publications

The total number of citations for an article (TC) reflects the authority and influence of this publication in the relevant field (Merigó *et al.*, 2015). Thus, Figure 5 shows the top 10 highly cited publications ordered by TC. The TCs of almost all the top 10 documents exceed 100. The article entitled "The effect of education and experience on self-employment success " that Robinson and Sexton (1994) wrote ranked first with 310 citations, which collected a great deal of data to discuss the role of education in the business success of entrepreneurs. It is followed by the article "Building an integrative model of small business growth" by (Wiklund *et al.*, 2009), which contributes to building a comprehensive model of small business growth by processing large amounts of data to analyze four subjective and objective factors. Finally, it is worth noting that the third-ranked article "The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms" by (Makridakis, 2017) was published in 2017 but is highly cited. This article explained the positive benefits of applying big data to enterprises. It concluded that those willing to take entrepreneurial risks and make extensive use of the Internet have a significant competitive advantage. Unfortunately, the TC does not give a complete picture of the impact of an article, which ignores the accumulation of numbers due to the time factor. Therefore, the number of citations per year (TC/Y) is introduced to give a more comprehensive picture to demonstrate the impact and authority of articles.

The top 10 articles ordered TC/Y are presented in Table 2. Multiple coauthors wrote more than 60% of the articles. Interestingly, almost all of the top 10 highly cited publications ordered by TC/Y were published from 2016 to 2018 in Period III (rapid development period) with the explosive growth of the BDIE literature. However, the articles published in Period I and Period II in Figure 5 are not listed in Table 2. Overall, these differences indicate that although some publications in Periods I and II have high citations due to time accumulation, they do not perform as well as some articles in Period III when TC/Y is counted, meaning that most articles in Period III have more significant influence and authority in this era.
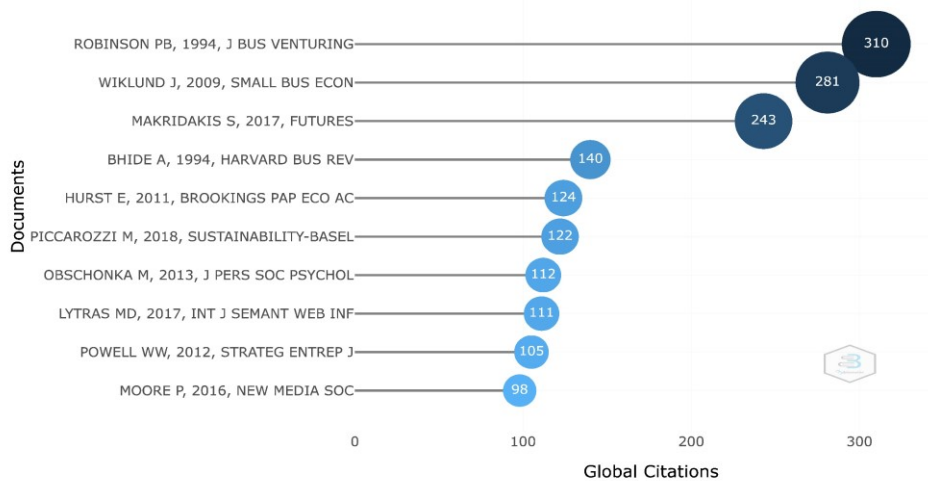

Figure 5. The Top 10 Highly cited Documents of the BDIE Ordered by TC

Table 2

The Top 10 Highly Cited Documents of the BDIE Ordered by TC/Y

Title	Year	TC	TC/Y	AN	Source
The forthcoming artificial intelligence (AI) revolution: Its impact on society and firms (Makridakis, 2017)	2017	240	48.00	1	FUTURES
Industry 4.0 in management studies: A systematic review of the literature (Piccarozzi *et al.*, 2018)	2018	122	30.50	3	SUSTAINABILITY
Big data and data analytics research: From metaphors to value space for collective wisdom in human decision making and smart machines (Lytras *et al.*, 2017)	2017	111	22.20	3	INTERNATIONAL JOURNAL ON SEMANTIC WEB AND INFORMATION SYSTEMS
Building an integrative model of small business growth (Wiklund *et al.*, 2009)	2009	281	21.62	3	SMALL BUSINESS ECONOMICS
The quantified self: What counts in the neoliberal workplace (Moore & Robinson, 2015)	2016	97	16.17	2	NEW MEDIA & SOCIETY
Publicising food: big data, precision agriculture, and co-experimental techniques of addition (Carolan, 2017)	2017	78	15.60	1	SOCIOLOGIA RURALIS
An exploration of contemporary organizational artifacts and routines in a sustainable excellence context (Carayannis *et al.*, 2017)	2017	75	15.00	5	JOURNAL OF KNOWLEDGE MANAGEMENT
The Impact of Digitalization on Business Models (Bouwman *et al.*, 2018)	2018	60	15.00	4	DIGITAL POLICY REGULATION AND GOVERNANCE
Towards a critique of cybernetic urbanism: the smart city and the society of control (Krivy, 2018)	2018	55	13.75	1	PLANNING THEORY
Integrating psychological approaches to entrepreneurship: the entrepreneurial personality system (EPS) (Obschonka & Stuetzer, 2017)	2017	62	12.40	2	SMALL BUSINESS ECONOMICS

Analysis of Countries/Regions and Authors

(1) At the Countries/Regions Level

With 63 countries/regions involved in the BDIE research, the performance of the top 20 countries/regions is presented in Figure 6, which is ordered by NP per country/region. Immediately, it can be seen that the first country/region is China with 151 articles, followed by the US and Italy with 135 and 69 papers. Respectively, when TC and AC indicators are introduced to measure the influence and authority of each country/region's articles. Surprisingly, China ranks first in the NP, but the AC of China is extremely low at four, even falling out of the top 20. Consequently, in future studies, China should pay more attention to improving the quality of articles while maintaining a high volume of publications.

Moreover, although countries/regions like Sweden and Canada only hover around the 20th in terms of NP, they are in the top 5 countries/regions in terms of AC, with 52.56 and 40.50, respectively. Similarly, Slovenia has only published nine articles but has a whopping 80 citations per article. Finally, it should be noted that the USA ranks second in NP and seventh in AC, showing a good performance in the enthusiasm of articles published and the quality of articles.
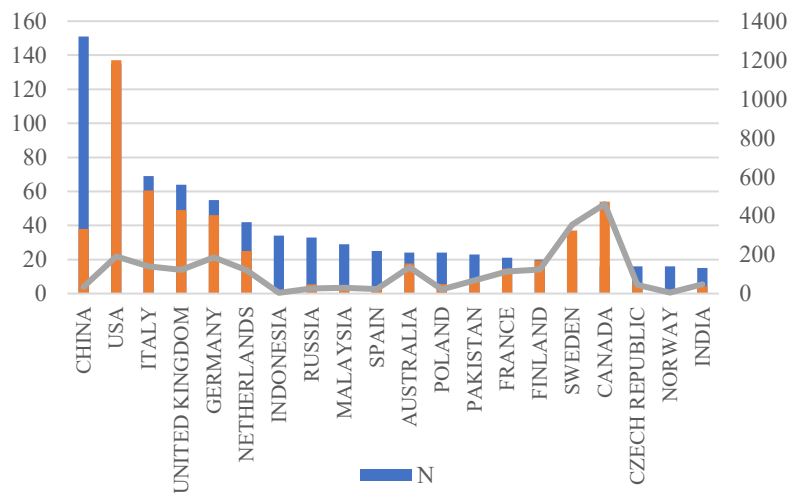

Figure 6. Top-20 Countries/Regions Ordered by NP

Table 3

Top-20 countries/regions ordered by AC

Ranked	Region	NP	TC	AC
1	SLOVENIA	9	240	80.00
2	ESTONIA	4	59	59.00
3	CANADA	16	473	52.56
4	SWEDEN	19	324	40.50
5	SLOVAKIA	12	119	29.75
6	BELGIUM	4	73	24.33
7	USA	135	1,200	21.82
8	GERMANY	55	403	21.21
9	ECUADOR	3	40	20.00
10	PERU	2	20	20.00
11	ITALY	69	530	16.06
12	AUSTRALIA	24	155	15.50
13	SERBIA	7	93	15.50
14	FINLAND	20	170	14.17
15	SWITZERLAND	6	14	14.00
16	UK	64	429	13.84
17	NETHERLANDS	42	219	13.69
18	FRANCE	21	129	12.90
19	SINGAPORE, 10 May 2001	4	11	11.00
20	IRAN	10	54	10.80

After a brief analysis of the individual countries/ regions, we need to show the international cooperation between them to explore the social structure in the BDIE research. Figure 7 illustrates the social network by showing the collaboration world map in BDIE research by setting the minimum edge to two in Bibliometrix. The color depth of the countries/regions on the map shows the number of publications, and the darker the color, the more productive the country/region is. They have a more positive attitude and investment in BDIE research. Moreover, the red line in the graph shows the cooperation between countries/regions. The more comprehensive the line, the more frequent the partnership of countries/regions. At first glance, China and the United States are prominent in the picture. As the leading publisher, China has worked with eight countries/regions, half of which are in Asia, such as Pakistan, Malaysia, and Singapore. Furthermore, the USA has cooperated with 14 countries/regions, becoming the country with the highest frequency of cross-border cooperation. Figure 7 clearly shows the high frequency of collaboration with European countries/regions, and the UK has the highest number of partnerships, ranking first with 159. The partnership between the United States and China is the strongest, indicating frequent exchanges and close cooperation in the BDIE research among scholars from the United States and China.
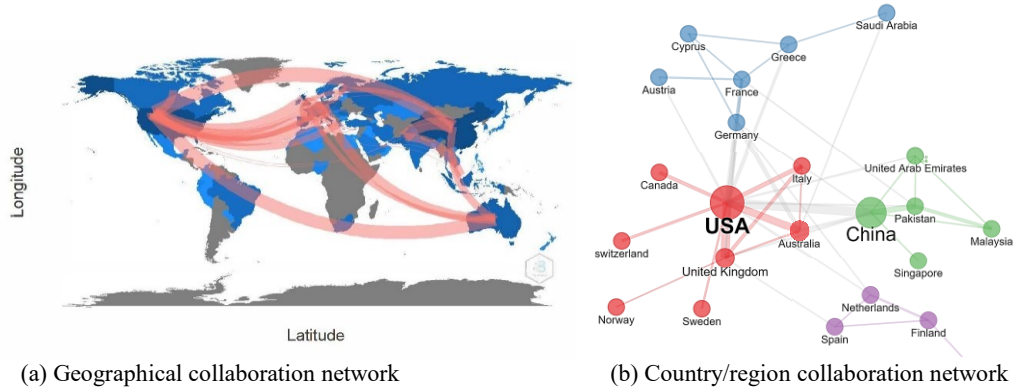

(a) Geographical collaboration network   (b) Country/region collaboration network

Figure 7. The Collaboration Networks of the Author's Country/Region

(2) At the Authors Level

Figure 8 with the timeline can help us analyze the authors' research in the temporal dimension with the support of Bibliometrix, which better illustrates the continuity and productivity of the scholar's research by showing the productive years of the scholars and the total number of citations received per year. The size and color depth of the circle indicates NP and TC/Y, respectively. The darker the circle, the higher the TC/Y, and the larger the circle, the higher the NP. Obschonka's research in this area lasted almost a decade from 2011 to 2020, and the high quality and quantity of results occurred in 2019. It produced four articles and received 19 citations per year. Silbereisen, RK, whose publications span nine years from 2011 to 2019, ranks second only to Obschonka, M, as one of the BDIE research pioneers. However, his research has a long gap with no article output for five years from 2014 to 2018, indicating poor continuity in the BDIE research. It is worth noting that some highly productive scholars, whose publications are concentrated in 2020, have a high potential and probability of contributing to BDIE research in the future (Purnomo A and Firdaus M).

As shown in Figure 9 (a), of 1,379 authors engaged in scientific research, 500 authors were selected to build the author collaboration network using the VOS viewer, and time was overlayed on the author network to present the final overlay visualization. According to the visualization principle of the VOS viewer, the node's size is proportional to the number of articles, and the darker the color, the earlier the author posts the article. As can be seen, Obschonka, M is in the center of the picture and ranks first with 13 articles, followed by Silbereisen, RK, who has far fewer articles than Obschonka, M, with just four articles. Furthermore, we can see that a more significant number of authors appeared between 2018 and 2020, indicating that the application of big data in innovative companies has attracted more and more attention in recent years.

Moreover, some highly productive authors have a stable coauthorship cluster and are grouped into ten research communities with different colors to warn us. There are ten communities of different colors, with scholars Obschonka, M, Anwar, Passiante, G, as the respective central points. These significant points are more active and collaborate more closely with other authors in BDIE research. The most extensive research community in Figure 9 (b) is purple. Five representative authors are influential, including Obschonka, M, Silbereisen, RK, Potter, J, Schmitt-Rodermund, E and Gosling, SD, and all are in Table 4.
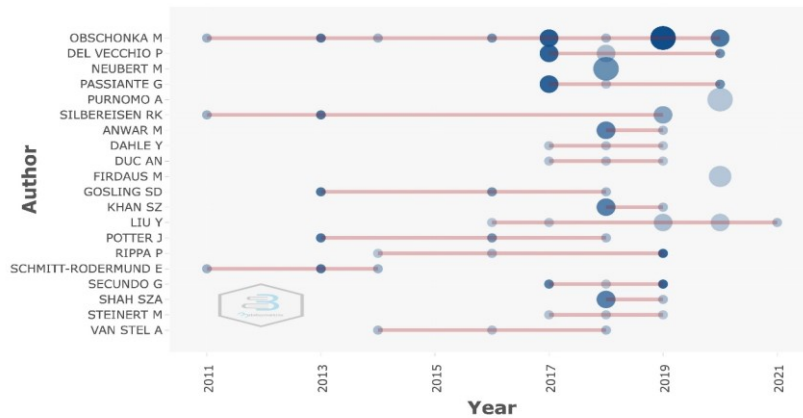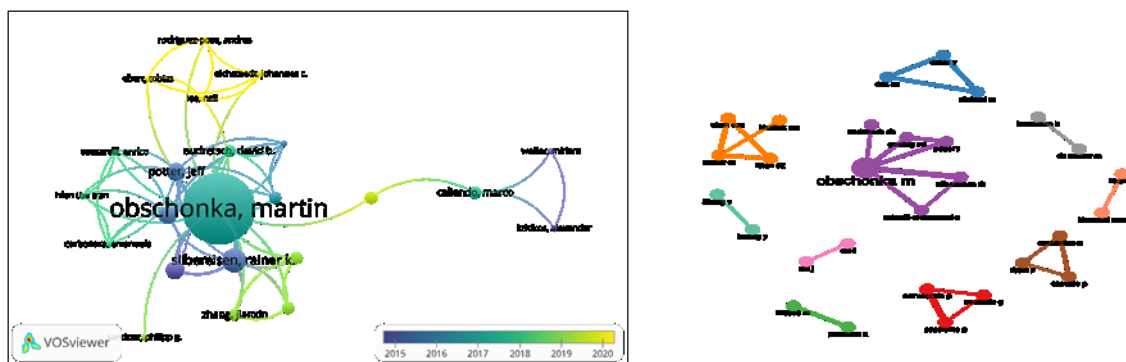

Figure 8. Production of the top 20 Authors Over Time Ordered by NP

Sources: Web of Science, powered by Bibliometrix

(a) Overlay of the collaborative network of authors.

(b) Collaboration network of authors.

Figure 9. Visualization Map of the Collaboration Network of Authors Concerning BDIE Research

Sources: Web of Science, powered by VOS viewer and Bibliometrix

Table 4

The H-index Orders the top 10 Authors

Author	h-index	g-index	m-index	TC	NP
OBSCHONKA M	9	13	0.818	373	13
SILBEREISEN RK	4	4	0.364	157	4
ANWAR M	3	4	None.	49	4
DEL VECCHIO P	3	4	0.600	96	4
GOSLING SD	3	3	0.333	162	3
KHAN SZ	3	3	0.750	48	3
PASSIANTE G	3	3	0.600	94	3
POTTER J	3	3	0.333	162	3
SCHMITT-RODERMUNDE	3	3	0.273	162	3
SHAH SZA	3	3	0.750	48	3

Source: Web of Science, analysis powered by Bibliometrix.

Keywords Analysis

(1) Co-Occurrence Analysis

Visual analysis of keywords, which are the high summarization and refinement of the article, shows the full picture of knowledge and indicates the direction of the current research (Bayramova *et al*., 2021). Furthermore, keyword analysis, such as co-occurrence and contingency matrix analysis, reflects the research directions and hot topics of articles in the field of BDIE (Xie *et al*., 2020). Before conducting the co-occurrence analysis, we present a treemap

in Figure 10 to get more basic information about keywords. In Figure 10, the treemap of keywords shows the top 50 high-frequency keywords in the BDIE research. Logically, the words "big data" and "entrepreneurship" appear 54 and 52 times at the top of the list, representing 7 % and 6 %, respectively, because they are part of our search strategy. In addition, "innovation", "performance", "management", and "impact" are other keywords with high frequency. In general, many scholars do BDIE research in conjunction with management and business, which plays an increasingly prominent role in supporting big data applications.
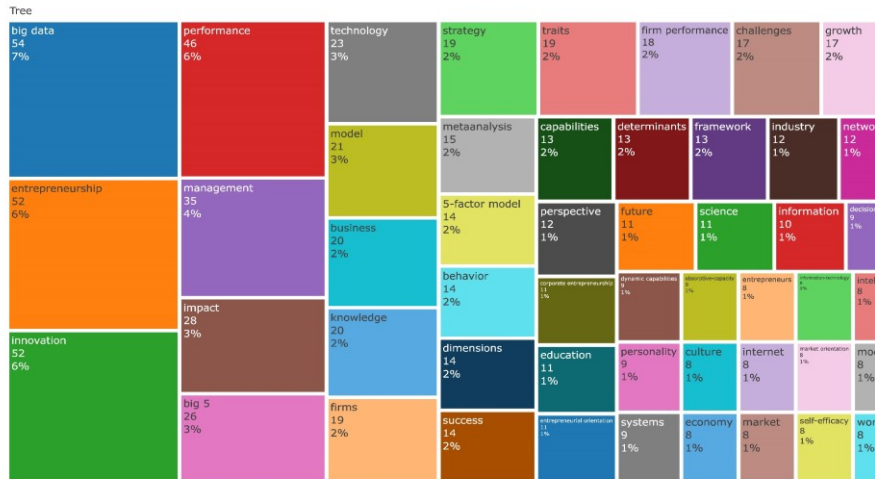

Figure 10. The Tree-Map of the 50 Highest-Frequency Keywords in the BDIE Research

Source: Web of Science, analysis powered by Bibliometrix

Keyword co-occurrence analysis can uncover more key knowledge and sort out the knowledge structure of BDIE research (Wang *et al*., 2020). To analyze the keyword co-occurrence, Bibliometrix software is employed to visualize the cooccurrence network of high-frequency keywords in the BDIE literature. Therefore, 50 high-frequency keywords form a cooccurrence network, as shown in Figure 11. Each node represents a keyword, and the node's size is proportional to the frequency. In addition, the distance between two nodes reflects their closeness, and the shorter the distance, the greater the degree of relationship between the two keywords.
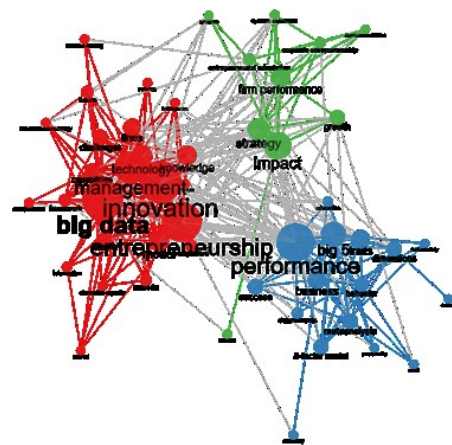
Furthermore, when related keywords are grouped, a cluster is formed. From this perspective, 50 items are divided into three clusters, that is, Cluster 1 with 24 keywords (red bubbles), Cluster 2 with 16 keywords (blue bubbles), and Cluster 3 with ten keywords (green bubbles). There are three different colored clusters with "Big data", "Performance", and "Impact" as the respective central point. They are more active and strongly associated with other keywords in the BDIE research. In detail, Cluster 1, being the most robust, mainly focuses on the innovation of big data technologies in enterprise management. Cluster 2 emphasizes how companies improve their performance by applying business models. Lastly, Cluster 3 explores corporate entrepreneurship and performance growth strategies.

Figure 11. Co-Occurrence Network of Keywords

Source: Web of Science, analysis powered by Bibliometrix

(2) Contingency Matrix

The contingency matrix is presented to visualize the joint distribution of two fields, showing the relevance between any pair of terms A_i and B_j in the sets A and B (Marvuglia *et al*., 2020). Furthermore, the contingency matrix is indicated by three colors: red, blue, and white, which represent different levels of relevance between the terms A_i and B_j. When many documents mention A_i also mention B_j, the cell is red, which means the most correlated. On the contrary, blue means anticorrelated when the few

documents mention A_i also mention B_j, while white cells do not feature any correlation, suggesting that joint mentions between A_i and B_j are neither more nor less than average. The right-hand side of the image has color-corresponding values. For example, a cell with a value of three means that the two fields are jointly mentioned 300 % more often than expected and vice versa.

The contingency matrix of keywords and countries/regions is shown in Figure 12, in which the number of terms is limited to 10. Our results show that highly relevant pairings include keywords "personality" and Germany, "big data analysis" and UK, and "entrepreneurship" and Russia. For example, when focusing on the keywords "big data" and "big data analysis" related to search strategy, we find that Germany pays the slightest attention to the research on big data. At the same time, Italy, China and the UK are very interested in that. Interestingly, although both keywords "big data" and "entrepreneurship" are included in the search strategy, the Italian, China, and the UK mentioned above also are anti-correlated in terms of entrepreneurship, indicating that the three countries seriously neglect the combination of entrepreneurship and big data. However, it is worth noting that the US performs better among the ten countries/regions, which strikes a balance between big data and entrepreneurial research.

In the same way, we show the contingency matrix of keywords and authors in Figure 13. Again, discernible highly positive relevance is observed between "SME's" and Anwar, M, "personality traits" and Anwar, M, "big data analysis" and Neubert, M, and "big data" and Del Vecchio, P. It is commendable that some experts have not only studied a particular subject in-depth but have focused on a variety of topics, such as Silbereisen, RK, and Obschonka, M. They focus on four topics "big five", "personality", "entrepreneurship" and "social media", becoming authors who covered the most research topics.
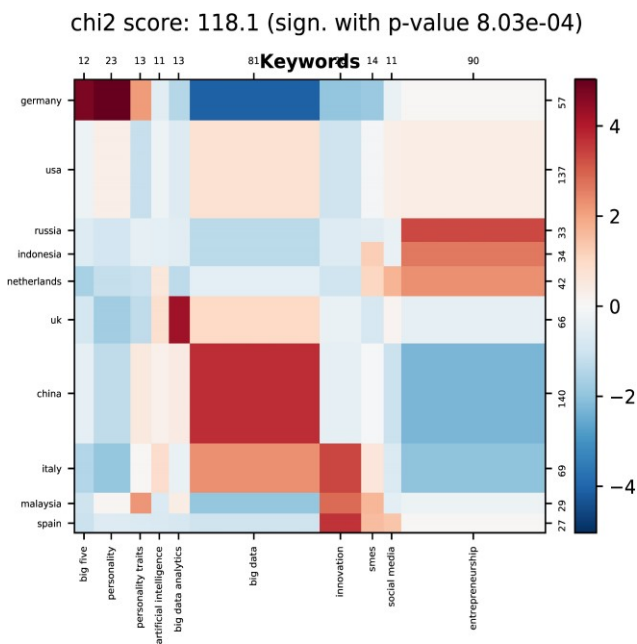
Figure 12. Contingency Matrix of Keywords and Countries/Regions

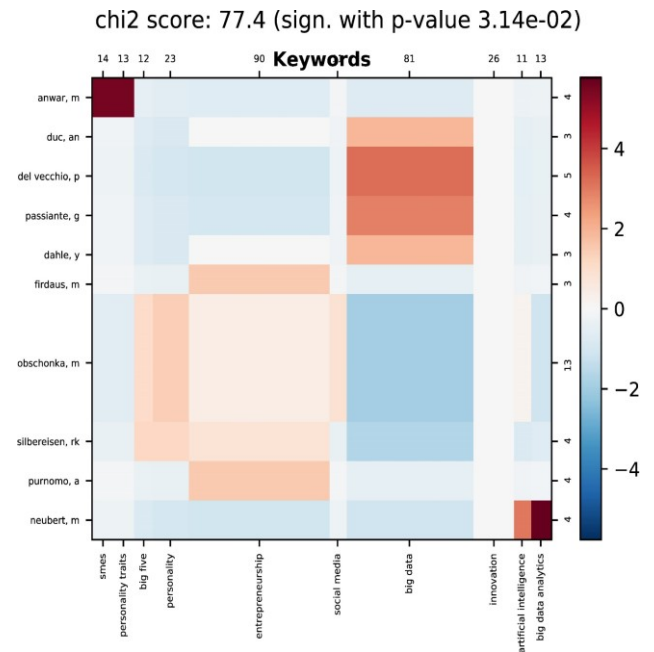*Source: Web of Science analysis powered by CorTexT Manager*

Figure 13. Contingency Matrix of Keywords and Authors

Source: Web of Science, analysis powered by CorTexT Manager

(3) Thematic Analysis and Evolution

As a research field developed over time, a thematic map composed of several specific themes is formed for scholars to distinguish and study. Therefore, Bibliometrix can provide a thematic map of the BDIE research (1993–2020) to identify themes based on cooccurrence analysis for keywords plus, which uses centrality as the horizontal coordinate (X-axis) and density as the vertical coordinate (Y-axis). Centrality measures the degree of interaction among clusters, while density measures the level of cohesion within clusters (Forliano *et al.*, 2021). The thematic map is segregated into four quadrants that are mother themes, niche themes, emerging and declining themes, and primary themes to showcase the unique traits of every theme. In this setting, the motor themes (the quadrant 1) have high centrality and density, indicating a bright future for development and a significant position in the BDIE research. Niche themes (the quadrant 2) have a good development, but they are easily underappreciated themes in BDIE research, which have high density and low centrality. In complete contrast to the motor themes in the first quadrant, disappearing or emerging themes are clustered in the quadrant 3, with not only low centrality but also low density. In reality, emerging and declining themes also represent slow development and low importance. Lastly, basic themes (the quadrant 4) with high centrality and low-density values represent some basic and long-established themes. What can be seen here is that each theme group is made up of several high-frequency keywords. Furthermore, the higher the frequency of keywords per topic, the correspondingly larger the size of the circle will be.

In this context, the following five topics in the BDIE research are divided into different regions. The theme with keywords "performance", "big", and "business" is located in

the first quadrant, suggesting that it has an important leading position and profound research significance. Moreover, the keywords "field", "big muddy", and "promise" form the niche themes that are located in the second quadrant, indicating that it is significant but isolated, therefore, sufficient interaction between groups is required to break this isolation. Finally, in the fourth quadrant, it is expected that the keywords "big data", "entrepreneurship", and "innovation" compose basic themes, suggesting that the application of big data in entrepreneurship and innovation is an ongoing topic of discussion.
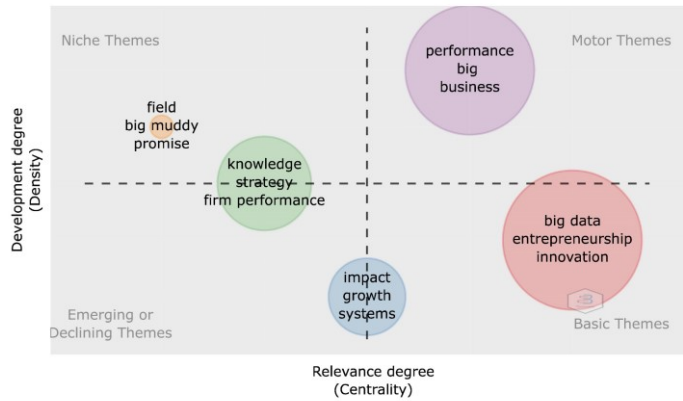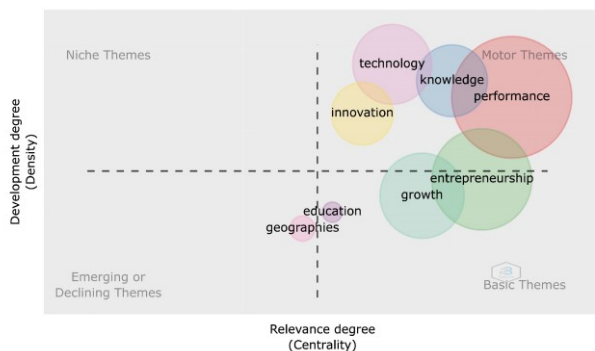

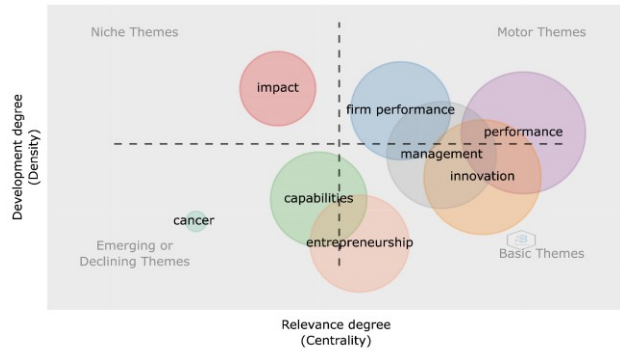**Figure 14.** Thematic Map of BDIE Research (1993–2020)

Thematic evolution analysis studies the changing rules, evolutionary relations, and evolutionary process of topics, which plays a vital role in revealing, grasping, and predicting the development trend of a domain (Xie *et al.*, 2020). Although a static thematic map through the data set from 1900–2020 is instructive and meaningful, more information is available when we consider the temporal dimension and analyze the evolution of the themes from a dynamic perspective (Qin *et al.*, 2022). Next, based on the number of publications, we set the cutting points as 2016, 2019, and 2020 for the dataset, dividing the data into four subperiods, and they are 1993–2016, 2017–2018, 2019–2020, and 2021–2021, respectively. As a result, Figure 15 presents different thematic maps for four time periods. Furthermore, the Sankey diagram, which visualizes the thematic evolution over time, contributes to analyzing the flow of different themes of the BDIE research, quantifying information such as thematic flow, thematic flow direction, and the transformation relationship (Soundararajan *et al.*, 2014). As a result, combining the analysis of Figures 15 and 18, we obtain the following significant findings.

The keyword "performance" dominated the motor themes (the quadrant 1) in the early studies from 1993 to

2016 and was one of the most frequent and rapidly developing themes. However, from 2017 to 2018, it started to develop slower and moved towards the basic themes (the quadrant 4), nevertheless, it remained a key theme in the BDIE field. Furthermore, between 2019 and 2020, the size of the circle decreased significantly, suggesting that the frequency of the keyword "performance" decreased, indicating a decrease in the popularity of the theme over time. Furthermore, the keyword "entrepreneur" occupied the quadrants 1 and 3 in 1993–2016, and as time passes, it gradually deteriorated and eventually never appeared again in 2019–2021 in Figures 15 (c) and (d). Although surprisingly, we found the reason for this in Figure 16, thematic evolution shows that "entrepreneur" evolved into three themes: innovation, model and market in 2017–2018, and "innovation" has become the basic theme, replacing "entrepreneur" as the main focus of scholars and the cornerstone of recent research. Finally, it is worth mentioning that the keyword "big data" did not appear in 1993–2018, but with the development of technology, it dominated the primary themes with the highest frequency in 2019–2020.
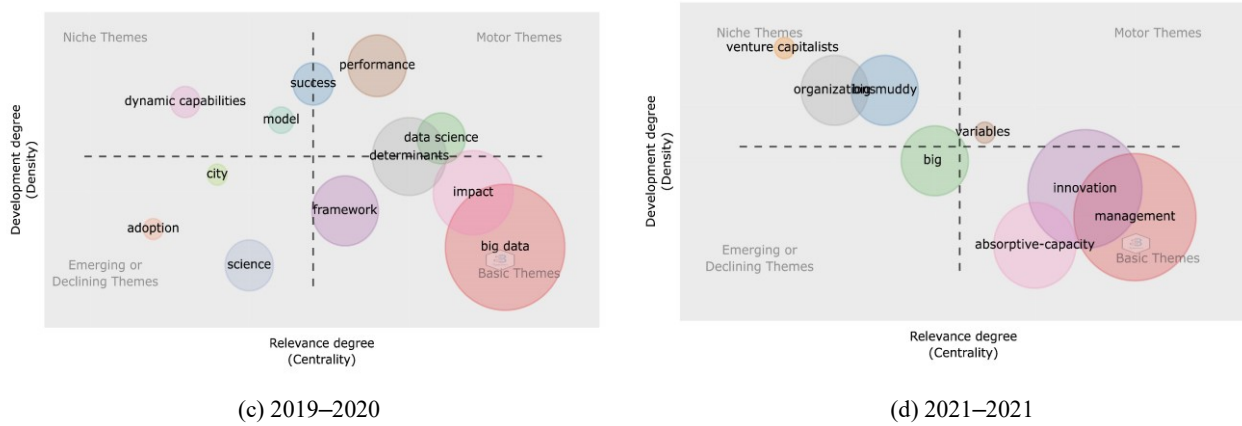

(a) 1993–2016

(b) 2017–2018

(c) 2019–2020   (d) 2021–2021

Figure 15. Thematic Map of the BDIE Research for Four Subperiods

Figure 16 shows the thematic evolution of BDIE research; As can be seen, scholars have maintained a high level of interest in some research topics from the beginning. For example, popular themes, such as "technology" and "entrepreneurship", are groundbreaking topics from the beginning, and they underwent thematic differentiation and branching in the next period (2017–2018). Similarly, we can see that innovation, which is based on four themes, has become the topic of the most attention in 2017–2018, and the three themes are "technology", "entrepreneurship" and "knowledge". Furthermore, the terms "capabilities" and "market" have evolved into "innovation" in a similar flow.

During this period, research themes have shifted from technology to market models and performance evaluation and had a greater focus on the use of innovative knowledge. In 2019–2020, with the development of technology and the emphasis on data technology, "big data" became the most popular theme, widely used in the area of decision-making in business activities. Furthermore, the emerging themes of "challenges" and "decision making" are attracting more and more attention, second only to "big data". Based on the theme "big data" flow trend, it is clear that big data will be used in more areas and evolve into more new themes in the future.
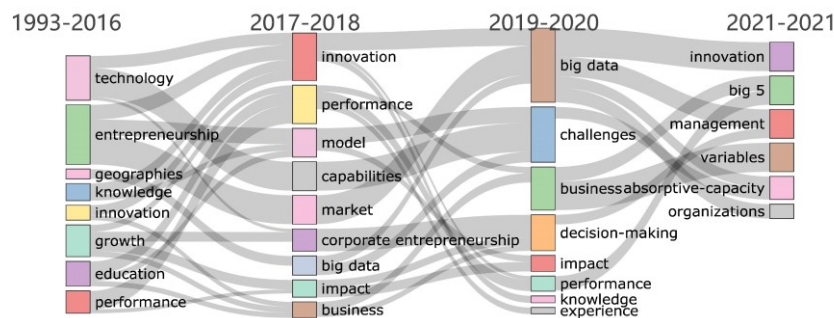

Figure 16. Thematic Evolution of BDIE Research (1993–2021)

Discussion

Popular Issues in the Current Study

BDIE research has evolved over the past 28 years since 1993 and has developed a complex framework. However, research is interdisciplinary and covers a wide range of research areas. Thus, we need to explore the primary research hotspots and provide scholars with new inspirational and innovative ideas. Thematic evolution and thematic maps vividly present the evolutionary process of topics and important research topics recently, demonstrating the breadth and popularity of the study over time. Therefore, we have summarized two topical issues in the current research of BDIE research based on the combined analysis of thematic evolution and thematic maps.

Big Data and Digital Entrepreneurship. Digital entrepreneurs should focus on strategy, leadership, management, and social media platforms (Wilk *et al*., 2021). Specifically, big data on social media platforms can be widely used by social entrepreneurs to assess the social impact of their businesses and promote the sustainability of social entrepreneurship globally (Zulkefly *et al*., 2021). In turn, entrepreneurship, which can drive the development of digital markets and digitalization, will have a significant impact on digital entrepreneurship (Purnomo *et al*., 2020). At the corporate level, intra-organizational entrepreneurship initiatives can help to promote the digital transformation of companies and the innovative development of organizational competencies (Ambos & Tatarinov, 2021). Therefore, digital disruptions can pose high risks for innovation management

and will affect steadily growing companies in the industry (Utoyo *et al*., 2020).

With the explosive growth of data and the increase in business applications, big data analysis has attracted a lot of attention from startups (Bahrami *et al*., 2021). When applying big data analysis capabilities (BDAC), BDAC is a key tool to improve business competitiveness in a highly dynamic market (Ciampi *et al*., 2021). However, the application of big data analysis techniques has been significantly influenced by factors such as the support of top management and the company's data-driven culture (Alsadi *et al*., 2021) Some computer algorithms are used in the management of university startups, and the use of big data technology can also be effective in improving their efficiency (Wang *et al*., 2020). At the level of entrepreneurial talent training, universities and enterprises focus on the cultivation of soft skills such as data collection and analysis and massive data manipulation (Luo, 2021). More companies focus on the potential demand for big data talents and build a model to train innovative entrepreneurial skills based on data mining technology to further optimize the innovative development of startups. (Deng & Wu, 2021; Yu & Zhang, 2021).

"Technology and business". The application of computer technology has made significant contributions to business and entrepreneurial research (Vitari & Raguseo, 2020). For example, Celbis (2021) has studied the use of five machine-learning techniques to explore factors related to the success and failure of potential entrepreneurs. Furthermore, analysis of accumulated data using programming programs contributes to exploring potential factors that influence the investment behavior of female entrepreneurs and their attitudes towards investing (Baig *et al*., 2021). On the one hand, datafication and AI-based machine labor have significantly impacted the labor market and are even considered new production factors (Wagner, 2020). On the other hand, human labor and AI need to find the right proportion and direction of application among companies (Popkova & Sergi, 2020).

On the other hand, however, technology promotes management skills. Mobile technology provides a way to develop digital skills, which allows people to actively participate in the digital economy and encourages the development of entrepreneurial skills (Venter & Daniels, 2020). However, Prufer and Prufer (2020) argued that entrepreneurial skills are of absolute importance, even more than digital skills for managers.

In summary, we provide a brief classification of the topical issues in the BDIE research, focusing on the following aspects. For example, what are the commercial applications of big data technology for startup enterprises, the relationship between datafication and HR within an organization, how big data technology can contribute to the development of entrepreneurial skills, and so on. However, related research is not limited to these aspects, and a growing number of scholars have explored and learned extensively in a wide range of fields and disciplines of BDIE research.

Enlightenment and Limitations

In the era of big data, the huge amount of information has a profound impact on the decision-making and development strategies of companies. Some big data analysis techniques, such as machine learning, data mining, and pattern recognition, have been widely used and developed in SMEs (Yismaw *et al*., 2021). Therefore, the application of big data in SMEs' entrepreneurship is a hot topic that has been widely discussed and has recently been a concern. At present, big data has powerful potential and promise for development, which is used in a wide range of disciplines including Business, Economics, Computer Science, Information, Education, Educational Research, etc. However, big data is still in its infancy and yet to be exploited, so it is vital to train business managers, technicians, and university students in the capabilities of big data analysis, which is conducive to stimulating entrepreneurship and developing new directions in business.

Digital companies are committed to developing business and commercial strategies based on big data and enhancing their technological competitiveness by providing customized services. However, the rapid development of digitization has led to massive amounts of data being mined and analyzed, which has also raised concerns about customer privacy and the security of company information leaks; in turn, it has influenced and contributed to the study of cyber security (Zhang *et al*., 2017). The development of the big data industry in the entrepreneurial sector faces enormous challenges, requiring good infrastructure, human and material resources, education to develop technical skills and entrepreneurial knowledge, and financial support to implement entrepreneurial strategies (Kim *et al*., 2016).

This paper employs bibliometric analysis to examine BDIE research through performance analysis and science mapping. While this approach is extensive and offers visualization and statistical analysis, it has limitations concerning data source, dataset, and search strategy. We collected data from the WoS database only, which could result in missing important literature. Also, the WoS does not include recent publications, leading to incomplete search results. Thus, future research should incorporate additional data sources like Google Scholar and Scopus. In terms of search strategy, we should adopt more precise keywords and constraints, which will increase the rigour of the study. Despite these limitations, this paper combines multiple dimensions and perspectives for an in-depth analysis of BDIE research and delves into the content of publications, which explores the evolutionary process of themes and relevance between terms through keyword analysis. To some extent, it compensates for the inability of bibliometric research to investigate content studies.

Conclusions

The rapid development of data-supported technologies has a significant impact on innovation and entrepreneurship. To gain a comprehensive understanding of the development of BDIE and to explore new research directions, we show the performance of BDIE publications, collaborative networks, and the evolution of themes, which help explore valuable research directions and evolutionary pathways of BDIE. We find that "digital entrepreneurship", "education of innovation and entrepreneurship" and "information technology in business" are future research directions and hot spots. Entrepreneurship drives the development of digital markets and digitalization, and there is a synergy between

entrepreneurship and the digital economy. However, to realize the big data industry in the field of entrepreneurship, it is extremely urgent to build an effective platform for entrepreneurial knowledge education. In addition, a digital entrepreneur faces the challenge of directing the coordinated activities of human resources, technology, and information more efficiently to achieve customer satisfaction. This paper provides scholars with new inspirational and innovative ideas and helps scholars understand the current state and limitations of BDIE research. This paper applies three bibliometric tools, i.e. Bibliometrix, VOS viewer, and CorTexT Manager, to perform a bibliometric analysis of BDIE research. The paper mainly comes from three aspects: performance analysis, analysis of countries/regions, authors, and keywords analysis, which focus on essential information and performance of publications, collaborative relationships, and exploration of themes, respectively. Finally, further discussions are made based on two directions: popular issues, enlightenment, and limitations. The main conclusions are summarized below.

(1) BDIE research shows a gradual increase in the number of publications and can be divided into three phases: 1993–2004 (startup period), 2005–2012 (initial development period), and 2013–2020 (rapid development period). Period III is the most prominent with an explosive increase in the number of articles; the same phenomenon is also found in the citation-year distribution. "Business" is the leading research direction, with 18 % among the ten subject categories of the BDIE research in WoS. Almost all of the top 10 highly cited publications ordered by AC were published in 2016–2018 in Period III, which means that publications in Period III have more significant influence and authority in this era.

(2) Up to now, 63 countries/regions participated in BDIE research, with China ranking first in terms of NP, followed by the USA and Italy. Furthermore, the collaboration between the USA and China is the strongest, reaching the highest frequency of association 11 times. At the author level, Obschonka, M's research in the BDIE lasted for a decade from 2011 to 2020, illustrating the continuity and productivity of his research. In the collaboration network, scholars are grouped into 10 research communities, and the team led by Obschonka, M, has good performance and a strong presence in the BDIE research.

(3) By analyzing the cooccurrence of keywords, there are three colored clusters with big data, performance, and impact as the respective central point. They show more robust and more active connections with other keywords. The contingency matrix shows that "personality" and Germany, "big data analysis" and the UK, and "entrepreneurship" and Russia are highly relevant pairings. Although Italy, China, and the UK strongly correlate with big data, they are anti-correlated with entrepreneurship. However, the USA strikes a balance between big data and entrepreneur research, performing better among the ten countries/ regions. Performance, big data, and business are the motor themes that occupy an important leading position and have profound research significance.

Acknowledgement

References

Akpor-Robaro, Mamuzo, & Oghen erobaro, M. (2012). The impact of socio-cultural environment on entrepreneurial emergence: an empirical analysis of Nigeria society. *Management Science & Eng*ineering. 6(4), 82–92. https://doi.org/10.3968/j.mse.1913035X20120604.3620

Alsadi, A. K., Alaskar, T. H., & Mezghani, K. (2021). Adoption of big data analytics in supply chain management: Combining organizational factors with supply chain connectivity. *International Journal of Information Systems and Supply Chain Management,* 14(2), 88–107. https://doi.org/10.4018/IJISSCM.2021040105

Ambos, T. C., & Tatarinov, K. (2021). Building responsible innovation in international organizations through intrapreneurship. *Journal of Management Studies*. https://doi.org/10.1111/joms.12738

Aria, M., & Cuccurullo, C. (2017). Bibliometrics: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Artz, J. M. (2013). Big data analytics: turning big data into big money. *Computing reviews*, 54(11), 659–660.

Bahrami, F., Kanaani, F., Turkina, E., Moin, M. S., & Shahbazi, M. (2021). Key challenges in big data startups: An exploratory study in Iran. *Iranian Journal of Management Studies*, 14(2), 273–289. https://doi.org/10.22059/Ijms.2020.303163.674082

Baier-Fuentes, H., Merigo, J. M., Ernesto Amoros, J., & Gaviria-Marin, M. (2019). International entrepreneurship: a bibliometric overview. *International Entrepreneurship and Management Journal*, 15(2), 385–429. https://doi.org/10.10 07/s11365-017-0487-y

Baig, U., Hussain, B. M., Davidaviciene, V., & Meidute-Kavaliauskiene, I. (2021). Exploring investment behavior of women entrepreneur: Some future directions. *International Journal of Financial Studies*, 9(2). https://doi.org/10.3390/ijfs9020020

Bayramova, A., Edwards, D. J., & Roberts, C. (2021). The role of blockchain technology in augmenting supply chain resilience to cybercrime. *Buildings*, 11(7), 283. https://doi.org/10.3390/buildings11070283

Bouwman, H., Nikou, S., Molina-Castillo, F. J., & de Reuver, M. (2018). The impact of digitalization on business models. *Digital Policy Regulation and Governance*, 20(2), 105–124. https://doi.org/10.1108/DPRG-07-2017-0039

Calic, G., & Ghasemaghaei, M. (2020). Big data for social benefits: Innovation as a mediator of the relationship between big data and corporate social performance. *Journal of Business Research*, 131, 391-401. https://doi.org/10.10 16/j.jbusres.2020.11.003

Carayannis, E. G., Grigoroudis, E., Del Giudice, M., Della Peruta, M. R., & Sindakis, S. (2017). An exploration of contemporary organizational artifacts and routines in a sustainable excellence context. *Journal of Knowledge Management*, 21(1), 35–56. https://doi.org/10.1108/JKM-10-2015-0366

Carolan, M. (2017). Publicising food: big data, precision agriculture, and co-experimental techniques of addition. *Sociologia Ruralis*, 57(2), 135–154. https://doi.org/10.1111/soru.12120

Celbis, M. G. (2021). A machine learning approach to rural entrepreneurship. *Papers in Regional Science*, 100(4), 1079–1104. https://doi.org/10.1111/pirs.12595

Chen, K., Jin, J., & Luo, J. (2021). Big consumer opinion data understanding for Kano categorization in new product development. *Journal of Ambient Intelligence and Humanized Computing*, 1–20. https://doi.org/10.1007/s12652-021-02985-5

Ciampi, F., Demi, S., Magrini, A., Marzi, G., & Papa, A. (2021). Exploring the impact of big data analytics capabilities on business model innovation: The mediating role of entrepreneurial orientation. *Journal of Business Research,* 123, 1–13. https://doi.org/10.1016/j.jbusres.2020.09.023

Ciampi, F., Marzi, G., Demi, S., & Faraoni, M. (2020). The big data-business strategy interconnection: a grand challenge for knowledge management. A review and future perspectives. *Journal of Knowledge Management.* https://doi.org/10. 1108/JKM-02-2020-0156

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. *Journal of Business Research*, 70. https://doi.org/10.1016/j.jbusres.2016.08.011

Deng, B. J., & Wu, J. (2021). The Cultivation of innovation and entrepreneurship skills and teaching strategies for college students from the Perspective of big data. *Arabian Journal for Science and Engineering*. https://doi.org/10.1007/s13 369-021-05893-0

Du, Q. Z., Li, J., Du, Y. Q., Wang, G. A., & Fan, W. G. (2021). Predicting crowdfunding project success based on backers' language preferences. *Journal of the Association for Information Science and Technology*. https://doi.org/10.10 02/asi.24530

Durana, P., Valaskova, K., Vagner, L., Zadnanova, S., Podhorska, I., & Siekelova, A. (2020). Disclosure of strategic managers' factotum: behavioral incentives of Innovative Business. *International Journal of Financial Studies*, 8(1), 17. https://doi.org/10.3390/ijfs8010017

Farooq, R., Rehman, S., Ashiq, M., Siddique, N., & Ahmad, S. (2021). Bibliometric analysis of coronavirus disease (COVID-19) literature published in Web of Science 2019-2020. *Journal of Family and Community Medicine*, 28(1), 1–7. https://doi.org/10.4103/jfcm.JFCM_332_20

Ferrati, F., & Muffatto, M. (2021). Entrepreneurial finance: emerging approaches using machine learning and big data. *Foundations and Trends in Entrepreneurship*, 17(3), 232–329. https://doi.org/10.1561/0300000099

Forliano, C., Bernardi, P. D., & Yahiaoui, D. (2021). Entrepreneurial universities: a bibliometric analysis within the business and management domains. *Technological Forecasting and Social Change*, 165, 120522. https://doi.org/10.10 16/j.techfore.2020.120522

Gao, P., Meng, F., Mata, M. N., Martins, J. M., Iqbal, S., Correia, A. B., Dantas, R. M., Waheed, A., Xavier Rita, J., & Farrukh, M. (2021). Trends and future research in electronic marketing: a bibliometric analysis of twenty years. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(5), 1667–1679. https://doi.org/10.3390/jtaer16050094

Gonzalez-Torres, T., Rodriguez-Sanchez, J.-L., Pelechano-Barahona, E., & Garcia-Muina, F. E. (2020). A systematic review of research on sustainability in mergers and acquisitions. *Sustainability*, 12(2), 513. https://doi.org/10.3 390/su12020513

Guerola-Navarro, V., Gil-Gomez, H., Oltra-Badenes, R., & Soto-Acosta, P. (2022). Customer relationship management and its impact on entrepreneurial marketing: a literature review. *International Entrepreneurship and Management Journal*. https://doi.org/10.1007/s11365-022-00800-x

Guleria, D., & Kaur, G. (2021). Bibliometric analysis of ecopreneurship using VOSviewer and RStudio Bibliometrics, 1989–2019. *Library Hi Tech*. https://doi.org/10.1108/LHT-09-2020-0218

Hao, Y., Tan, Q., & Li, Z. (2021). Research on the new application of information technology in the model of innovation and entrepreneurship education based on big data. 2021 2nd International Conference on *Information Science and Education*. https://doi.org/10.1109/ICISE-IE53922.2021.00060

Kim, H., Choi, M., Jeon, B., & Kim, H. (2016). A study on the big data business model for the entrepreneurial ecosystem of the creative economy. Advances in Parallel and Distributed Computing and *Ubiquitous Services*, 368, 158–190. https://doi.org/10.1007/978-981-10-0068-3_24

Kittichotsatsawat, Y., Jangkrajarng, V., & Tippayawong, K. Y. (2021). Enhancing coffee supply chain towards sustainable growth with big data and modern agricultural technologies. *Sustainability*, 13(8). https://doi.org/10.3390/su13084593

Krivy, M. (2018). Towards a critique of cybernetic urbanism: the smart city and the society of control. *Planning Theory*, 17(1), 8–30. https://doi.org/10.1177/1473095216645631

Lei, Y. A., A, C., & Xy, B. (2019). Pricing and carbon emission reduction decisions considering fairness concern in the big data era. *Procedia CIRP*, 83, 743-747. https://doi.org/10.1016/j.procir.2019.04.325

Linnenluecke, M. K., Marrone, M., & Singh, A. K. (2020). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management*, 45(2), 175–194. https://doi.org/10.1177/0312896219877678

Lipych, L., Khilukha, O., & Kushnir, M. (2021). Interdependence between entrepreneurship, innovation and competencies. *Intellect XXI.* https://doi.org/10.32782/2415-8801/2021-1.8

Luo, L. S. (2021). Research on the cultivation mode of application-oriented e-commerce talents under the background of smart new retail. *2021 2nd International Conference on E-Commerce and Internet Technology*, 146–150. https://doi.org/10. 1109/Ecit52743.2021.00040

Lytras, M. D., Raghavan, V., & Damiani, E. (2017). Big data and data analytics research: from metaphors to value space for collective wisdom in human decision making and smart machines. International *Journal on Semantic Web and Information Systems*, 13(1), 1–10. https://doi.org/10.4018/IJSWIS.2017010101

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: its impact on society and firms. *Futures*, 90, 46-60. https://doi.org/10.1016/j.futures.2017.03.006

Manogaran, G., Thota, C., & Kumar, M. V. (2016). MetaCloudDataStorage architecture for big data security in cloud computing. *Procedia Computer Science*, 87, 128–133. https://doi.org/10.1016/j.procs.2016.05.138

Manyika, J., Chui, M., Brown, B., Bughin, J., & Byers, A. H. (2011). Big data: the next frontier for innovation, competition, and productivity.

Mariani, M. M., & Nambisan, S. (2021). Innovation analytics and digital innovation experimentation: the rise of research-driven online review platforms. *Technological Forecasting and Social Change*, 172, 121009. https://doi.org/10.10 16/j.techfore.2021.121009

Marvuglia, A., Havinga, L., Heidrich, O., Fonseca, J., Gaitani, N., & Reckien, D. (2020). Advances and challenges in assessing urban sustainability: an advanced bibliometric review. *Renewable and Sustainable Energy Reviews,* 124, 109788. https://doi.org/10.1016/j.rser.2020.109788

MaryAnne, M., & Gobble. (2015). Big Data: the next big thing in innovation. *Research-Technology Management*, 56(1), 64–67. https://doi.org/10.5437/08956308X5601005

Meng, X., & Chan, A. H. S. (2021). Current states and future trends in safety research of construction personnel: a quantitative analysis based on social network approach. *International Journal of Environmental Research and Public Health*, 18(3), 883. https://doi.org/10.3390/ijerph18030883

Merigo, J., Mas-Tur, A., Roig-Tierno, N., & Ribeiro-Soriano, D. (2015). A bibliometric overview of the Journal of Business Research between 1973 and 2014. *Journal of Business Research*, 68(12), 2645–2653. https://doi.org/10.10 16/j.jbusres.2015.04.006

Moore, P., & Robinson, A. (2015). The quantified self: What counts in the neoliberal workplace. *New Media & Society*, 18(11). https://doi.org/10.1177/1461444815604328

Moral-Munoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: an up-to-date review. *Profesional De La Informacion*, 29(1), 290103. https://doi.org/10.3145/epi.2020.ene.03

Obschonka, M. (2017). The quest for the entrepreneurial culture: psychological big data in entrepreneurship research. *Current Opinion in Behavioral Sciences*, 18, 69–74. https://doi.org/10.1016/j.cobeha.2017.07.014

Obschonka, M., & Stuetzer, M. (2017). Integrating psychological approaches to entrepreneurship: the Entrepreneurial Personality System (EPS). *Small Business Econ*omics, 49(1), 203–231. https://doi.org/10.1007/s11187-016-9821-y

Pan, Y., Shi, H., & Niu, G. (2021). The choice of financing mode for serial entrepreneurs in the big data. *2021 International Conference on Applications and Techniques in Cyber Intelligence*, Cham. https://doi.org/10.1007/978-3-030-79200-8_124

Park, Y. E. (2021). Developing a COVID-19 crisis management strategy using news media and social media in big data analytics. *Social Science Computer Review*, 089443932110073. https://doi.org/10.1177/08944393211007314

Piccarozzi, M., Aquilani, B., & Gatti, C. (2018). Industry 4.0 in management studies: a systematic literature review. *Sustainability*, 10(10). https://doi.org/10.3390/su10103821

Pogrebna, G. (2015). Big data, brand loyalty, and business models: accounting for imprecision and noise in consumer preferences. *Wmg Service Systems Research Group Working Paper*.

Popkova, E. G., & Sergi, B. S. (2020). Human capital and AI in industry 4.0. Convergence and divergence in social entrepreneurship in Russia. *Journal of Intellectual Capital*, 21(4), 565–581. https://doi.org/10.1108/JIC-09-2019-0224

Prufer, J., & Prufer, P. (2020). Data science for entrepreneurship research: studying demand dynamics for entrepreneurial skills in the Netherlands. *Small Business Economics*, 55(3), 651–672. https://doi.org/10.1007/s11187-019-00208-y

Purnomo, A., Firdaus, M., Sutiksno, D. U., Latukismo, T. H., & Rachmahani, H. (2020). A study of digital market status using the bibliometric approach during four decades. *2020 International Conference on Information Management and Technology* (ICIMTech), 458–463. https://doi.org/10.1109/ICIMTech50083.2020.9211201

Qin, Y., Xu, Z. S., Wang, X. X., & Skare, M. (2022). Green energy adoption and its determinants: a bibliometric analysis. *Renewable and Sustainable Energy Reviews*, 153, 111780. https://doi.org/10.1016/j.rser.2021.111780

Qin, Y., Xu, Z. S., Wang, X. X., Skare, M., & Porada-Rochon, M. (2021). Financial cycles in the economy and in economic research: A case study in China. *Technological and Economic Development of Economy*, 27(5), 1250–1279. https://doi.org/10.3846/tede.2021.15439

Robinson, P. B., & Sexton, E. A. (1994). The effect of education and experience on self-employment success. *Journal of Business Venturing*, 9(2), 141–156. https://doi.org/10.1016/0883-9026(94)90006-X

Soundararajan, K., Ho, H. K., & Su, B. (2014). Sankey diagram framework for energy and exergy flows. *Applied Energy*, 136, 1035–1042. https://doi.org/10.1016/j.apenergy.2014.08.070

Urban, B., & Mutendadzamera, K. (2021). Social capital leading to innovation: understanding moderating effects of the environment in the Zimbabwean small and medium enterprise context. *Journal of Enterprising Communities-People and Places in the Global Economy*. https://doi.org/10.1108/JEC-01-2021-0010

Utoyo, I., Fontana, A., & Satrya, A. (2020). The role of entrepreneurial leadership and configuring core innovation capabilities to enhance innovation performance in a disruptive environment. *International Journal of Innovation Management*, 24(6). https://doi.org/10.1142/S1363919620500607

Vaneck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

Venter, I. M., & Daniels, A. D. (2020). Towards bridging the digital divide: The complexities of the South African story. 14th International Technology, *Education and Development Conference* (Inted2020), 3250–3256. https://doi.org/10.21125/inted.2020.0937

Vitari, C., & Raguseo, E. (2020). Big data analytics business value and firm performance: linking with environmental context. *International Journal of Production Research*, 58(18), 5456–5476. https://doi.org/10.1080/00207543.2019.1660822

Wagner, D. N. (2020). Economic patterns in a world with artificial intelligence. Evolutionary and *Institutional Economics Review*, 17(1), 111–131. https://doi.org/10.1007/s40844-019-00157-x

Walter, C. E., Valente, T., Polonia, D. F., Au-Yong-Olivera, M., & Veloso, C. M. (2021). Big data, European data strategy and innovation: a systematic review of the literature. *Quality-Access to Success*, 22(184), 16–20. https://doi.org/10.47750/QAS/22.184.02

Wan, W. H., & Liu, L. J. (2021). Intrapreneurship in the digital era: driven by big data and human resource management? *Chinese Management Studies*, 15(4), 843–875. https://doi.org/10.1108/CMS-07-2020-0282

Wang, C., Dong, Y. Z., Xia, Y. J., Li, G. X., Martinez, O. S., & Crespo, R. G. (2020). Management and entrepreneurship management mechanism of college students based on support vector machine algorithm. *Computational Intelligence*, 1–13. https://doi. org/10.1111/coin.12430

Wang, X. X., Chang, Y. R., Xu, Z. S., Wang, Z. D., Kadirkamanathan, V. (2020). 50th anniversary of International Journal of Systems Science: A comprehensive bibliometric analysis. *International Journal of Systems Science*, https://doi.org/10.1080/00207721.2020.1862937

Wang, X. X., Xu, Z. S., & Skare, M. (2020). A bibliometric analysis of Economic Research-Ekonomska Istrazivanja (2007-2019). *Economic Research-Ekonomska Istrazivanja*, 33(1), 865–886. https://doi.org/10.1080/1331677X.2020.1737558

Wang, X. X., Xu, Z. S., Su, S. F., & Zhou, W. (2021). A comprehensive bibliometric analysis of uncertain group decision making from 1980 to 2019. *Information Sciences*, 547, 328–353. https://doi.org/10.1016/j.ins.2020.08.036

Wang, Y., & Ali, Z. (2021). Exploring big data use to predict supply chain effectiveness in Chinese organizations: a moderated mediated model link. *Asia Pacific Business Review*. https://doi.org/10.1080/13602381.2021.1920704

Wiklund, J., Patzelt, H., & Shepherd, D. A. (2009). Building an integrative model of small business growth. Small Business Economics, 32(4), 351–374. https://doi.org/10.1007/s11187-007-9084-8

Wilk, V., Cripps, H., Capatina, A., Micu, A., & Micu, A. E. (2021). The state of digital entrepreneurship: a big data Leximancer analysis of social media activity. *International Entrepreneurship and Management Journal,* 1899–1916. https://doi.org/10.1007/s11365-020-00729-z

Xie, H. L., Zhang, Y. W., Wu, Z. L., & Lv, T. G. (2020). A bibliometric analysis on land degradation: current status, development, and future directions. *Land*, 9(1), 28. https://doi.org/10.3390/land9010028

Xie, T. (2021). Research on the development of innovation and entrepreneurship education in universities under the background of big data. *the 2021 2nd International Conference on Big Data and Informatization Education*. https://doi.org/10.1109/ICBDIE52740.2021.00016

Yismaw, M. B., Tesfaye, Z. T., & Bhagavathula, A. S. (2021). Assessment of pharmacy students' satisfaction towards pharmacotherapy lectures delivered at the University of Gondar, Gondar, Ethiopia. *Education Research International*, 2021, 5601773. https://doi.org/10.1155/2021/5601773

Yu, D. J., Xu, Z. S., Pedrycz, W., & Wang, W. R. (2017). Information Sciences 1968-2016: a retrospective analysis with text mining and bibliometric. *Information Sciences*, 418, 619–634. https://doi.org/10.1016/j.ins.2017.08.031

Yu, X., & Zhang, B. G. (2021). Innovation strategy of cultivating innovative enterprise talents for young entrepreneurs under higher education. *Frontiers in Psychology*, 12. https://doi.org/10.3389/fpsyg.2021.693576

Zeng, J. (2018). Fostering path of ecological sustainable entrepreneurship within big data network system. *International Entrepreneurship and Management Journal*, 14(1), 79–95. https://doi.org/10.1007/s11365-017-0466-3

Zhang, Y., Huang, Y., Porter, A. L., Zhang, G., & Lu, J. (2017). Discovering interactions in big data research: a learning-enhanced bibliometric study. 2017 Portland International Conference on *Management of Engineering and Technology (Picmet)*, 1–12. https://doi.org/10.23919/PICMET.2017.8125292

Zheng, J. L., Qiao, H., Zhu, X. M., & Wang, S. Y. (2021). Knowledge-driven business model innovation through the introduction of equity investment: evidence from China's primary market. *Journal of Knowledge Management*, 25(1), 251–268. https://doi.org/10.1108/JKM-02-2020-0158

Zhou, C. J., & Wang, D. X. (2021). A risk assessment algorithm for college student entrepreneurship based on big data analysis. *Complexity*, 2021. https://doi.org/10.1155/2021/6359296

Zhou, R., Ming, L., & Tao, L. (2014). Characterizing the efficiency of data deduplication for big data storage management. Workload Characterization (IISWC), *2013 IEEE International Symposium on*, 98–108. https://doi.org/10.1109/IISWC.2013.6704674

Zulkefly, N. A., Ghani, N. A., Hamid, S., Ahmad, M., & Gupta, B. B. (2021). Harness the global impact of big data in nurturing social entrepreneurship: a systematic literature review. *Journal of Global Information Management*, 29(6). https://doi.org/10.4018/JGIM.20211101.oa18

Zurita, R. T., Milian, M. J. R., & Diaz, P. L. (2021). Digitization practices implemented in companies from human resources departments: critical analysis of the discourse. *Prisma Social* (32), 498–525.

Authors' Biographies

Anran Xiao is currently working toward a master degree in Business School, Sichuan University, China. Her research interests include intelligent decision-making, total quality management, multi-attribute decision-making, and supply chain management. Her research results have been published in the International Journal of Retail & Distribution Management, International Journal of Energy Research, and International Journal of Environmental Research and Public Health. [ORCID: 0000-0002-0638-2287]

Yong Qin is currently working toward a Ph.D. degree in Business School, Sichuan University, China. His research interests include decision analysis, tourism management, information fusion and bibliometrics. His research works have been published in Renewable and Sustainable Energy Reviews, Journal of Cleaner Production, International Entrepreneurship and Management Journal, Technological and Economic Development of Economy, International Journal of Fuzzy Systems, Economic Analysis and Policy, etc. [ORCID: 0000-0002-4966-7899].

Zeshui Xu is currently a Professor with the Business School, Sichuan University, Chengdu. His current research interests include Decision-making theory and methodology, optimization algorithms, information fusion, and big data analytics. He was a Distinguished Young Scholar of the National Natural Science Foundation of China and the Chang Jiang Scholar of the Ministry of Education of China. He is currently the Senior Editor of IEEE Access, and the Associate Editor of IEEE Transactions on Cybernetics, IEEE Transactions on Fuzzy Systems, Information Sciences, Artificial Intelligence Review, Cognitive Computation, Applied Intelligence, Journal of the Operational Research, Fuzzy Optimization and Decision Making, etc.

Marinko Škare is a Professor of Economics, Economic Research Journal Editor in Chief, Vice Rector for Research and Art and Member of Editorial Board of several international journals. He served as Assistant Dean for Education, Assistant Dean for International Cooperation, Faculty of Economics & Tourism, Pula, Main and Team Researcher on several scientific projects, Former Dean of the Faculty of Economics & Tourism, Pula and Former Vice President for International Cooperation, Juraj Dobrila University of Pula. He has published several books and many scientific papers about economic growth, welfare economics and poverty. [ORCID: 0000-0001-6426-3692].