Big Data in the Innovation Process – A Bibliometric Analysis and Discussion

Zornitsa Yordanova(✉)

University of National and World Economy, Sofia, Bulgaria
zornitsayordanova@unwe.bg

Abstract. Companies are already using big data to develop innovations and improve efficiency and productivity in many business processes. The innovation process is one of the most important processes at the company level and beyond, as it impacts the company's overall performance and potential. Innovation is a data-intensive process, so the development of new technologies, especially big data, has a significant impact on innovation. The purpose of this research is to examine how big data is currently influencing the innovation process. The current scope of big data in innovation processes is defined by a literature analysis of 110 publications indexed in the Web of Science, Social Science Citation Index. This represents any article that uses both "big data" and "innovation process" in the title, abstract, and author words. In addition to state-of-the-art analysis, we identify topics that have not yet been explored to advance future research in this area and in some clusters, allowing further depth and analysis of the topic.

Keywords: Innovation Process · Big Data · Innovation Management · Technology Management · Bibliometric Analysis

1 Introduction

Big data as a tool for innovation continues to grow in popularity [1], but it is imperative to examine how the nature and steps of the innovation process change as a result of its inclusion in its development and application, as they become components of the process and the knowledge management within enterprises [2]. Research on innovation's curves and transformation is urgently needed in light of its significant influence on overall innovations it brings as a main systematic means of innovating across teams, organizations, and industries, as well as the growing research into technologies involved in its digital transformation [3, 4]. Recently, many studies have provided evidence on the scope and level of big data (BD) use for improving firm performance [5] but still, an overall perspective and mapping of the whole literature for BD use in innovation management is missing in regard to the innovation process as a separate theoretical sub-filed of the innovation management science. Thousands of comprehensive research has already revealed the role of BD in co-innovation, open and collaborative innovation [6], service innovation [7], and new product development [8] as well as in some

M. Papadaki et al. (Eds.): EMCIS 2022, LNBIP 464, pp. 122–133, 2023.
https://doi.org/10.1007/978-3-031-30694-5_10

particular economic fields such as agriculture [9], hospitality sector [7], banking [10], etc. The majority of the literature indicates that BD has a positive impact on innovation performance [11], but it has yet to be explored how innovation patterns are changing. For addressing firms' goals and challenges, data has become increasingly critical [12]. The concept of business data management as a superior level of data management can open up the littoral capabilities of business data management towards external sources of data that firms could incorporate into their internal processes in order to gain operational efficiency, boost the efficiency of processes, support decision making [9] and accelerate go-to-market strategies as well as increase customer satisfaction [13]. Since BD has been hailed as the next big thing in innovation [14], it has not only been used for product development but has also been embedded into some companies' innovation processes to increase the sustainability and performance of any kind of innovative development.

The innovation process is among the essential business process in organizations that directly influence a firm's profits [15]. Innovation for firms is often assessed as the most important outcome of business processes and critical for a firm's performance [16]. All of these factors place the internal innovation process at the center of a firm's innovation capabilities and the entire company's ability to operate successfully.

As part of the firm's innovation process, big data incorporation and utilization are inevitable, especially when research indicates that BD can positively influence employee creativity [17] as well as company performance and decision-making [18]. Nevertheless, the literature does not provide a clear picture of the current state of the art or evidence regarding the successful implementation of BD. As a promising research area, proven by the increasing number of studies on the topic and generally on the topic of the use of BD in the business process of firms for improving the quality and results of them, this study also aims at identifying the areas for further research for the innovation processes in particular. Furthermore, to gain a clear understanding of the development of this research area, it is critical to examine the current literature in its entirety. According to Sheng, Amankwah-Amoah, and Wang [19], big data in business processes and big data in innovation management represent emerging research areas in which multidisciplinary expertise is required. Bibliometric analysis is a useful method to analyse cross-discipline, emerging, or rapidly changing research areas impacting diverse areas [20, 21]. In our case, we examine a firm's innovation process and how big data technology is influencing it.

The results from the study give insights into the current state of the art on the problem of BD usage in the innovation process and give directions for further research.

2 Theoretical Background

2.1 Firm's Innovation Process

The firm's innovation process or internal innovation process (FIP) has attracted the interest of researchers since the 1950s and it is still amongst the hottest sub-topics in management science. In general, the innovation process is a sequence of activities and strategic vision related to the development and commercialization of innovative outcomes from organizations to the audience/market or for internal needs. Its evolution has gone through different concepts and theories starting from more general and industry-based therefore internal-to-firm models. Rothwell [22] provided a systematization into five categories: The Technology Push Theory – in the 1950s, The Market Pull Theory – in the 1960s, The Coupling Innovation Process Theory – in the 1970s–1980s, The Functional Integrated Innovation Process Theory [23] – 1980s, The Systems Integration and Networking Innovation Process Theory - 1990s.

Later on, more FIP have been introduced and followed as principles by firms such as Open Innovation [24], User Innovation [25], Lean Startup framework [26], chain-linked models [27], the Information assurance of innovation process [22], Innovation process based on continuous improvement [28], etc. with diverse kinds of extensions and customizations. Opening the innovation process not only towards different stakeholders but also to different concepts [29] incl. Including emerging technologies such as artificial intelligence [30], technology acceptance models, or BD [31] have clearly defined benefits in several case studies within the scientific literature. Nowadays, FIP remains under-researched as modern techniques, technologies, and design tools for adjusting the innovation processes are frequently applied [32].

However, a common understanding of the innovation process is missing in the literature, as it is quite complex and specific to the firm's sector and approach [33]. Nevertheless, among the most important aspects of the innovation process is identified to be the need for gathering information from and transmitting information to several external information areas within and outside of the organization, called also innovation systems [34]. BD provides a huge amount of data, incl. Internal and external to the firm that may benefit the organizational innovation process [35]. BD use in the innovation process as a managerial approach influencing the process for developing not only technological innovations directly related to BD, but any kind of products or processes is still under-researched, which motivates this article.

2.2 Big Data in Innovation Management and Innovation Process

Innovation management is considered a comprehensive and general scientific field for innovation management in organizations. The innovation management and specifically firm's innovation process theories contain numerous sub-streams that are generally clustered into different themes and directions for research and practice. Here in this section, we are revealing some of these sub-streams of theories, which are already discussed and researched in the context of big data as part of the innovation process, incl. The use of big data, data analytics, decision-making through big data, big data for generating users' ideas, big data to support innovation commercialization, etc. The digital transformation of firms and even of entire industrial sectors calls for the introduction of new forms of human-machine collaboration through the increased use of big data and cognitive systems require completely new approaches [36]. As Trabucchi and Buganza scholars [37] suggested, scholars are constantly providing different strategies and methods to help firms grasp and understand the added value immersed in their data to foster innovation and improve the efficiency of existing processes. To enhance organizational learning in the innovation process, improved use of existing and new information and its better assimilation become imperative [2]. Furthermore, big data benefits not on the innovation process but also firm productivity in general [38].

3 Research Design

The research design is organized following the main principles of bibliometric analysis as this method can extract relevant information from a large number of publications and elicit the information for answers to the research questions in this study.

3.1 Data Selection

The scope of this research was defined by conducting a Boolean search in the Web of Science (WoS) database for extracting high-quality publications regarding the innovation process and big data that are mutually discussed and analyzed. The formula used was the following:

> ALL: BIG DATA (Topic) and ALL: INNOVATION PROCESS (Topic), refined by: document types: article; language: English; Social Sciences Citation Index (SSCI).

After conducting the search, 110 publications met the search criteria. The details about the scope are provided in Table 1 and presented visually in Table 1.

Table 1. Data selection and scope for bibliometric analysis

Timespan	1998:2022
Sources (Journals, Books, etc.)	77
Documents	110
Annual growth rate %	12.25
Average citations per document	26.63
Authors	352
Co-Authors per Doc	3.3
International co-authorships %	34.55%

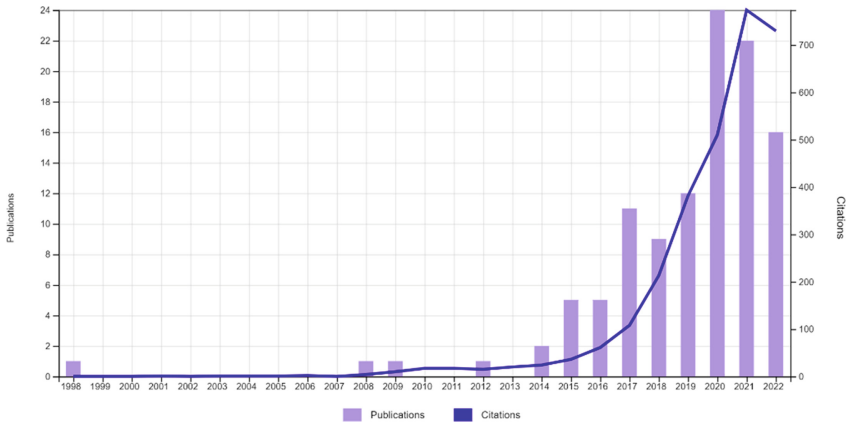

Fig. 1. Annual scientific production on BD in the innovation process

Figure 1 shows the exponential development of research fields that drive research. Even if the research does not focus on a specific research question, its value and contribution can be attributed to mapping the current state of the art in the interdisciplinary scientific field of big data in the innovation process and to one of the identified areas.

3.2 Bibliometric Analysis

Bibliometric analysis was applied as a systematically proven type of research by many scholars and is currently considered one of the most effective scientific methods for understanding the research field from a historical, holistic, and interdisciplinary perspective. Bibliometric analysis facilitates the mapping of the current research as well as identifies knowledge gaps, streams of research already done, and authors' information, and recognizes further research agenda. Bibliometric analysis is an effectively used

method to explore the emergence of the domain of digitalization and innovation [39] and has thacanthushe research and forecast future research trends [25].

In this study, we applied the following bibliometric analysis to address the core topic of digitalization of the firm's innovation process in the particular field of using BD:

- Co-word analysis
- Top-tier Journals
- Themati

Thematicose of bibliometric analysis, we used R Studio and its package Biblioshiny.

4 Results and Discussion: Big Data in the Innovation Process

The first results show the most commonly met keywords used by authors, the most influential authors, and the respective journals publishing such research. Amongst the journals, three out of four have a more sustainable focus. Another interesting insight is the combination of artificial intelligence along with BD (Fig. 2).

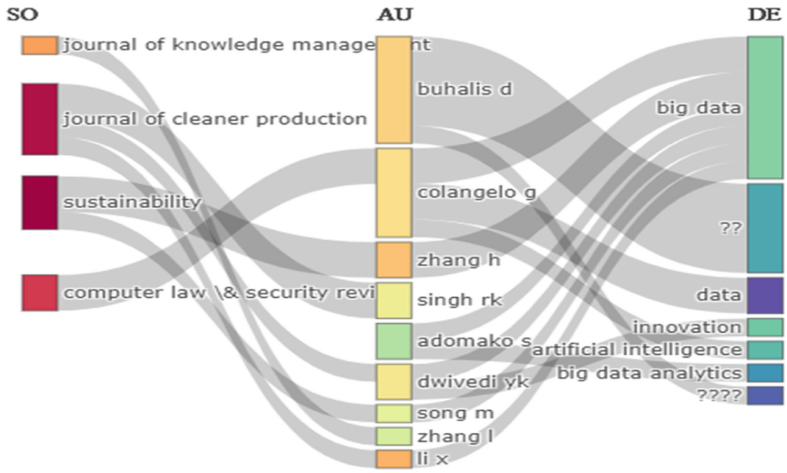

Fig. 2. Three plot analyses on sources, authors, and author keywords

Figure 3 reveals the most relevant sources for such research: Sustainability and Journal of cleaner production, which both are eco-oriented.

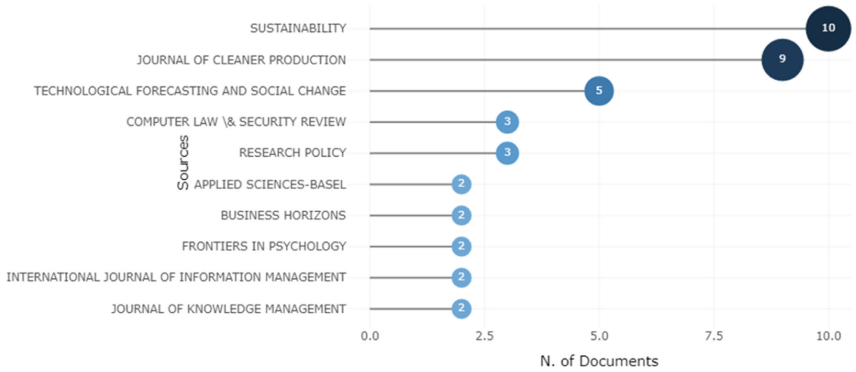

Fig. 3. Most relevant sources (journals)

Figure 4 presents the most influential studies on the matter. The one on first place has already been cited by 480 other studies.

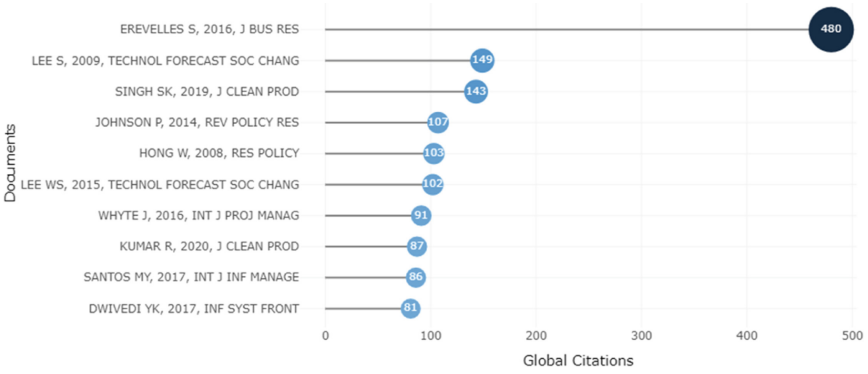

Fig. 4. Most influential research

The bigram co-word analysis reveals sectors that BD is researched and applied in – the healthcare sector and hospitality. It also gives clues on the type of companies that usually apply BD in the innovation process, which are the corporates. When it comes to business functions, BD is used in the supply chain and social media performance. Other technologies that are also researched in combination with BD are AI and could computing. The results are presented in Fig. 5.

The cumulative co-word analysis demonstrates the huge increase in the number of studies discussing BD and the innovation process after 2017. Mostly, it has been in the context of analytics, firm performance, and management (Fig. 6).

Trending topics are presented in Fig. 7 using bigrams word analysis. The trending words are arranged in time periods interesting observation brings the insight that in the last two years, BD has been mainly discussed in the context of digital transformation,

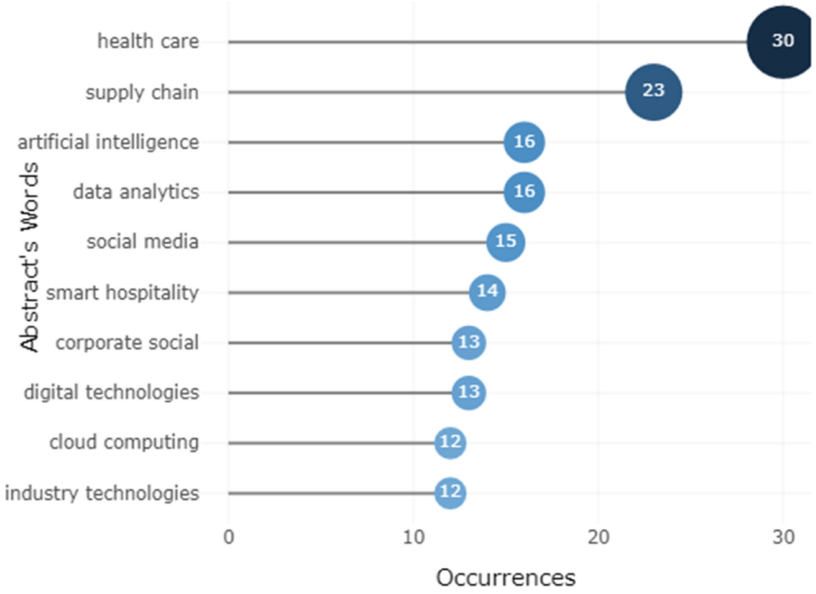

Fig. 5. Co-word analysis on abstracts (bigrams)

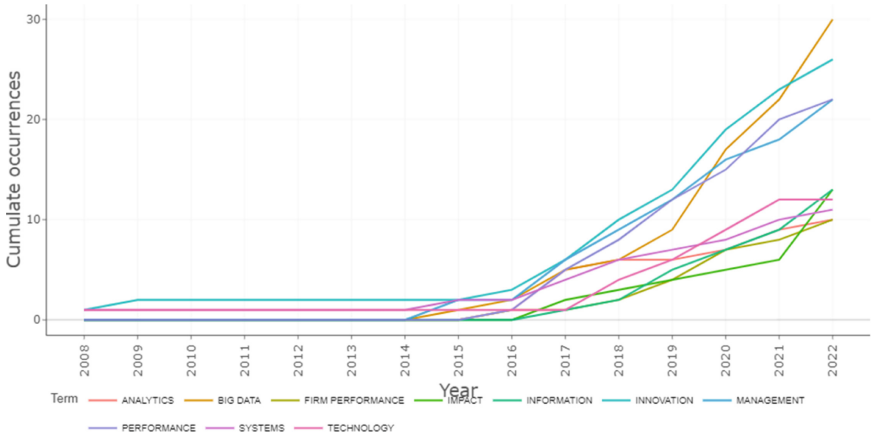

Fig. 6. Cumulative co-word analysis (unigrams)

digital innovation and, supply-chain management. The topic emerged as more a sustainable related application back in 2012 associated with social responsibility and innovation climate. In the last years, it was already incorporated into the innovation process along with other technologies bringing automation efficiency.

The thematic mapping shows different aspects of developing the topic in the literature. The basic topics are well researched and these are related to machine learning

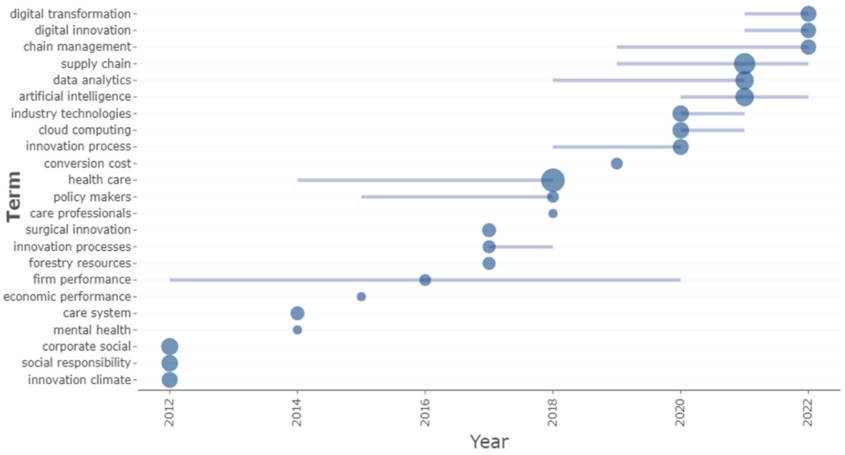

Fig. 7. Co-word analysis, bigrams over the years

(in combination with BD), innovation processes, and technological and digital innovations. The motor themes are those, which have gotten the strongest research interest recently. These are innovation systems, cloud computing, digital technologies, digital transformation, AI, and analytics.

For further research in the next years when the growing trends of publications and experimentation of the use of BD in the innovation process will continue to exponentially increase, we recommend a systematic analysis to be conducted for clear patterns and concrete use to be revealed. In such a systematic analysis, border publications and out-of-scope science fields should be extracted and several inclusion and exclusion criteria should be applied to organizing best practices for managerial purposes. This study can be a motivation for further research in either of the two scientific fields as a new common denominator.

5 Conclusion

This study provides a bibliometric analysis of the scientific literature indexed in the Web of Science on BD use in the firm's innovation process. Generally, we can summarize the current application of BD in this specific firm's activity into the following groups:

- Big data for decision-making in innovation
- Big data in innovation management and innovation process
- Big data to encourage innovation
- Big data to facilitate innovation
- Big data to scale up innovation (incl. Innovation performance)
- Big data as a methodological approach in innovation studies

Further research should focus on motor themes and still under-researched combinatory applications of BD along with other technologies such as cloud computing and

AI. From the innovation management domain, the main theoretical knowledge areas on which the scoped studies were focused on aro-innovation, service innovation, innovation process, product innovation, and process innovation.

Acknowledgment. This work was supported by the UNWE Research Programme (Research Grant No 9/2021).

References

1. Ghasemaghaei, M., Calic, G.: Assessing the impact of big data on firm innovation performance: big data is not always better data. J. Bus. Res. **108**, 147–162 (2020). https://doi.org/10.1016/j.jbusres.2019.09.062
2. Johannessen, J.A., Olsen, B., Olaisen, J.: Aspects of innovation theory based on knowledge-management. Int. J. Inf. Manage. **19**(2), 121–139 (1999). https://doi.org/10.1016/s0268-4012(99)00004-3
3. Marion, T.J., Fixson, S.K.: The transformation of the innovation process: how digital tools are changing work, collaboration, and organizations in new product development. J. Prod. Innov. Manag. **38**(1), 192–215 (2020). https://doi.org/10.1111/jpim.12547
4. Appio, F.P., Frattini, F., Petruzzelli, A.M., Neirotti, P.: Digital transformation and innovation management: a synthesis of existing research and an agenda for future studies. J. Prod. Innov. Manag. **38**(1), 4–20 (2021). https://doi.org/10.1111/jpim.12562
5. Cappa, F., Oriani, R., Peruffo, E., McCarthy, I.: Big data for creating and capturing value in the digitalized environment: unpacking the effects of volume, variety, and veracity on firm performance. J. Prod. Innov. Manag. **38**(1), 49–67 (2020). https://doi.org/10.1111/jpim.12545
6. Bresciani, S., Ciampi, F., Meli, F., Ferraris, A.: Using big data for co-innovation processes: mapping the field of data-driven innovation, proposing theoretical developments and, providing a research agenda. Int. J. Inf. Manage. **60**, 102347 (2021). https://doi.org/10.1016/j.ijinfomgt.2021.102347
7. Shamim, S., Yang, Y., Zia, N.U., Shah, M.H.: Big data management capabilities in the hospitality sector: Service innovation and customer generated online quality ratings. Comput. Hum. Behav. **121**, 106777 (2021). https://doi.org/10.1016/j.chb.2021.106777
8. Duan, Y., Cao, G., Edwards, J.S.: Understanding the impact of business analytics on innovation. Eur. J. Oper. Res. **281**(3), 673–686 (2020). https://doi.org/10.1016/j.ejor.2018.06.021
9. Su, Y., Wang, X.: Innovation of agricultural economic management in the process of constructing smart agriculture by big data. Sustain. Comput.: Inform. Syst. **31**, 100579 (2021). https://doi.org/10.1016/j.suscom.2021.100579
10. Hassani, H., Huang, X., Silva, E.: Digitalization big data mining in banking. Big Data Cognitive Comput. **2**(3), 18 (2018). https://doi.org/10.3390/bdcc2030018
11. Lin, R., Xie, Z., Hao, Y., Wang, J.: Improving high-tech enterprise innovation in big data environment: a combinative view of internal and external governance. Int. J. Inf. Manage. **50**, 575–585 (2020). https://doi.org/10.1016/j.ijinfomgt.2018.11.009
12. Chen, H., Chiang, R.H.L., Storey, V.C.: Business intelligence and analytics: from big data to big impact. MIS Q. **36**(4), 1165 (2012). https://doi.org/10.2307/41703503
13. Akter, S., Wamba, S.F., Gunasekaran, A., Dubey, R., Childe, S.J.: How to improve firm performance using big data analytics capability and business strategy alignment? Int. J. Prod. Econ. **182**, 113–131 (2016). https://doi.org/10.1016/j.ijpe.2016.08.018

14. Gobble, M.M.: Big data: the next big thing in innovation. Res. Technol. Manag. **56**(1), 64–67 (2013). https://doi.org/10.5437/08956308x5601005
15. Galanakis, K.: Innovation process. Make sense using systems thinking. Technovation **26**(11), 1222–1232 (2006). https://doi.org/10.1016/j.technovation.2005.07.002
16. Hitt, M., Hoskisson, R., Johnson, R.A., Moesel, D.D.: The market for corporate control and firm innovation. Acad. Manag. J. **39**(5), 1084–1119 (1996). https://doi.org/10.2307/256993
17. Dammak, H., Dkhil, A., Cherifi, A., Gardoni, M.: Enterprise content management systems: a graphical approach to improve the creativity during ideation sessions—case study of an innovation competition "24 h of innovation." Int. J. Interact. Des. Manuf. (IJIDeM) **14**(3), 939–953 (2020). https://doi.org/10.1007/s12008-020-00691-8
18. Delen, D., Zolbanin, H.M.: The analytics paradigm in business research. J. Bus. Res. **90**, 186–195 (2018). https://doi.org/10.1016/j.jbusres.2018.05.013
19. Sheng, J., Amankwah-Amoah, J., Wang, X.: A multidisciplinary perspective of big data in management research. Int. J. Prod. Econ. **191**, 97–112 (2017). https://doi.org/10.1016/j.ijpe.2017.06.006
20. Dhir, S., Dhir, S.: Diversification: literature review and issues. Strateg. Chang. **24**(6), 569–588 (2015). https://doi.org/10.1002/jsc.2042
21. Dhir, S., Ongsakul, V., Ahmed, Z.U., Rajan, R.: Integration of knowledge and enhancing competitiveness: a case of acquisition of Zain by Bharti Airtel. J. Bus. Res. **119**, 674–684 (2020). https://doi.org/10.1016/j.jbusres.2019.02.056
22. Rothwell, R.: Industrial Innovation: Success, Strategy, Trends. The Handbook of Industrial Innovation. Edward Elgar, Aldershot (1994)
23. Imai, K., Nonaka, I., Fakeuchi, H.: Managing the new product development. In: Clark, K., Hayes, F. (eds.) The Uneasy Alliance. Harvard Business School Press (1985)
24. Chesbrough, H.W.: Open Innovation: The New Imperative for Creating and Profiting from Technology (First Trade Paper ed.). Harvard Business Review Press (2003).
25. von Hippel, E.: The dominant role of users in the scientific instrument innovation process. Res. Policy **5**(3), 212–239 (1976). https://doi.org/10.1016/0048-7333(76)90028-7
26. Ries, E.: The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses (Illustrated ed.). Currency (2011)
27. Kline, S.J., Rosenberg, N.: An overview of innovation. In: Rosenberg, N. (ed.) Studies on Science and The Innovation Process Selected Works of Nathan Rosenberg, Chapter 9, pp. 173–203, World Scientific Publishing Co. Pte. Ltd. (2009)
28. Tidd, J., Bessant, J.R.: Managing Innovation: Integrating Technological, Market and Organizational Change (7th edn.). Wiley (2020)
29. Gassmann, O.: Opening up the innovation process: towards an agenda. R D Manag. **36**(3), 223–228 (2006). https://doi.org/10.1111/j.1467-9310.2006.00437.x
30. Ohlberg, K.H., Salmeron, J.L.: Proposal-based innovation: a new approach to opening up the innovation process. In: Glauner, P., Plugmann, P. (eds.) Innovative Technologies for Market Leadership. FBF, pp. 201–231. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-41309-5_14
31. Brous, P., Janssen, M., Herder, P.: The dual effects of the Internet of Things (IoT): a systematic review of the benefits and risks of IoT adoption by organizations. Int. J. Inf. Manage. **51**, 101952 (2020). https://doi.org/10.1016/j.ijinfomgt.2019.05.008
32. Lendel, V., Hittmár, T., Siantová, E.: Management of innovation processes in company. Procedia Econ. Finan. **23**, 861–866 (2015). https://doi.org/10.1016/s2212-5671(15)00382-2
33. Becheikh, N., Landry, R., Amara, N.: Lessons from innovation empirical studies in the manufacturing sector: a systematic review of the literature from 1993–2003. Technovation **26**(5–6), 644–664 (2006). https://doi.org/10.1016/j.technovation.2005.06.016
34. Tushman, M.L.: Special boundary roles in the innovation process. Adm. Sci. Q. **22**(4), 587 (1977). https://doi.org/10.2307/2392402

35. Lozada, N., Arias-Pérez, J., Perdomo-Charry, G.: Big data analytics capability and co-innovation: an empirical study. Heliyon **5**(10), e02541 (2019). https://doi.org/10.1016/j.heliyon.2019.e02541
36. Vocke, C., Constantinescu, C., Popescu, D.: Status quo and quo vadis: creativity techniques and innovative methods for generating extended innovation processes. Procedia CIRP **91**, 39–42 (2020). https://doi.org/10.1016/j.procir.2020.02.148
37. Trabucchi, D., Buganza, T.: Data-driven innovation: switching the perspective on Big Data. Eur. J. Innov. Manag. **22**(1), 23–40 (2019). https://doi.org/10.1108/ejim-01-2018-0017
38. Wu, L., Lou, B., Hitt, L.M.: Data analytics supports decentralized innovation. SSRN Electron. J. (2019) https://doi.org/10.2139/ssrn.3351982
39. Zhang, X., et al.: A bibliometric analysis of digital innovation from 1998 to 2016. J. Manag. Scie. Eng. **2**(2), 95–115 (2017). https://doi.org/10.3724/SP.J.1383.202005