

Drivers and Loci of Big Data Innovation: A Review, Synthesis, and Future Research Directions

Nebojša Radojević¹ and Erich Heumüller²

¹IU International University of Applied Sciences, Schweriner Str. 5, 01067 Dresden, Germany | nebojsa.radojevic@iu.org

²Baden-Wuerttemberg Cooperative State University, Center for Digital Transformation, Paulinenstr. 50, 70178 Stuttgart, Germany | erich.heumueller@dhw-stuttgart.de

Abstract

Despite a drastic increase in the amount of data collected and stored, the exact role that big data plays in innovation remains a subject of a diverging debate. This paper hence employs a systematic literature review and qualitative analysis techniques to identify, analyze and synthesize extant conceptual contributions on how big data drives the innovation process, where this process takes place, and which theoretical perspectives have shaped the research on big data innovation so far. Our findings reveal that big data innovation primarily takes place cooperatively, in the form of open innovation or within innovation networks, and that it is simultaneously driven by a variety of drivers. However, purposeful applications of big data and analytics for innovating in response to existing but unsatisfied market needs (“market pull”) remain comparatively under-researched. Concomitantly, the theoretical lenses of “traditional” innovation management have largely framed the research on big data innovation so far. We coherently integrate these findings within an overarching framework and suggest an extensive future research agenda.

Keywords: Big data, Innovation, Literature review.

Cite paper as: Radojević, N., Heumüller, E., (2023). Drivers and Loci of Big Data Innovation: A Review, Synthesis, and Future Research Directions, *Journal of Innovation Management*, 11(2), 42-70.; DOI: https://doi.org/10.24840/2183-0606_011.002_0003

1 Introduction

The diffusion of digital technologies has dramatically increased the amount of data collected and stored by organizations. In this context, big data in its narrower sense refers to datasets of scope and volume that would require vast amounts of storage and time to analyze with traditional systems and that conventional databases could not control or examine (Ward and Barker, 2013; Baig et al., 2019). In the broader sense, the term “big data” encapsulates both data itself and big data analytics (Nunan and Di Domenico, 2017). Extant literature typically focuses on the characteristics of big data (volume, velocity, variety, valence, veracity, variability, and vagueness), different subsets thereof (“3Vs” or “5Vs” of big data), approaches for analysis (e.g., business analytics, machine learning, data mining), and various aspects of *big data innovation*, i.e. *big data-based value creation channeled through innovation in products, services, business processes, or business models* (Chen et al., 2013; Demchenko et al., 2013; Chen and Zhang, 2014; Del Vecchio et al., 2018; Saggi and Jain, 2018).

However, the exact role that big data plays in innovation is a subject of a diverging debate, particularly so regarding drivers, loci, and applicable theoretical foundations. Stating that “information is the oil of the 21st century [whereas] analytics is the combustion engine”, Peter

Sondergaard from Gartner Research (cited in Yablonsky, 2019, p. 16) basically limits this role to being a digital commodity that drives innovation by its mere abundance and need for related technologies. According to Ma and Zhang (2021), big data has caused excessive enthusiasm although related technologies and theories have only reached the conceptual stage so far, rather than the maturity necessary for their practical application. In line with this, a relative majority of 5.700 manufacturers surveyed in 2019 consider big data to have low importance for innovation in products and business processes, while an absolute majority believe that the buying and/or selling of big data (open innovation; West and Bogers, 2014) will play a minor role for business activities (Gradeck et al., 2019). In contrast, according to Scuotto et al. (2017), big data improves both inter- and intra-organizational innovation processes, while Del Vecchio et al. (2018) and Trabucchi et al. (2018) assert that big data represents an emerging opportunity for open innovation. More generally, Blackburn et al. (2017) see the potential of big data to disrupt innovation management, in which it will hence play a crucial role in the future.

This divergence obscures the links between big data innovation and established theories of (innovation) management where they exist and fails to reliably identify where they do not, which prevents further research from being purposefully pursued, appropriately absorbed, and practically applied. Consequently, this paper identifies, consolidates, evaluates, and synthesizes the extant literature on how big data drives the innovation process at the firm (i.e., organizational) level, where this process takes place, and which theoretical perspectives towards it have been taken so far. By doing so, we theoretically contribute to the field with a comprehensive framework that coherently integrates both the individual theoretical lenses previously taken towards big data innovation and our findings regarding its drivers and loci. Along the elements of the same framework, we then suggest an extensive agenda for future research.

To the best of our knowledge, no study has systematically reviewed the literature on these aspects of big data innovation so far, although some recent papers (Buganza et al., 2020; Ciampi et al., 2020; Trabucchi and Buganza, 2019, 2022) address research questions partly related to ours. However, Buganza et al. (2020) elaborate on the big data-based evolution of services, which are only one form of the innovation outcome (the other being products, business processes, and business models; OECD and Eurostat, 2018). Ciampi et al. (2020) entirely focus on the relationship between big data and business strategy, while Trabucchi and Buganza (2019) propose an approach to innovation stimulated exclusively by big data in its narrower sense. Finally, Trabucchi and Buganza (2022) review the literature on multi-sided platforms, which is only one of the common business model patterns (Osterwalder and Pigneur, 2010). In contrast, our paper remains neutral regarding both the form of innovation outcome and business model pattern, distinguishes between big data in the narrower and broader sense, and sees business strategy as one of several potential determinants of big data innovation at the firm level (cf. Crossan and Apaydin, 2010).

This paper continues as follows. Section 2 briefly reviews major theoretical models of the innovation processes to summarize common drivers and loci of innovation in general. Based on these, we derive drivers and loci specific to big data innovation, which we later employ to define a supporting structure for our data analysis. Section 3 describes the literature review method used in our study. Section 4 performs a descriptive analysis of the final article sample. Section 5 presents the results of our literature review along the individual drivers of big data innovation, while considering its loci and theoretical foundations. Section 6 synthesizes our findings, while Section 7 outlines future research directions identified by this study. The final Section 8 highlights our contributions and concludes.

2 Theoretical Background

Innovation is a two-dimensional phenomenon that may signify both a process and the outcome of the process and that may occur at several levels, from a single firm to the whole world (OECD and Eurostat, 2018). At the firm (or, more generally, organizational) level, the innovation process is characterized by sub-dimensions such as the driver (or stimulus) and locus (Crossan and Apaydin, 2010). Different generations of models of the innovation process connect these sub-dimensions in a particular way by assuming a certain stimulus to innovate from which the process would flow in a specific direction (Rothwell, 1994; Hobday, 2005; Tidd et al., 2005). Drawing on Schumpeter's concept of creative destruction, the first generation or *technology-push models* conceptualize innovation as a sequential process driven by new technology that linearly flows from fundamental research via applied research, product development and production to marketing. Ever since the first generation, later models have challenged the respective dominant view by re-positioning the driver or locus, or by revealing that the direction of the innovation process deviates or inverts. In this vein, second generation models (*market-pull*; Schmookler, 1966) invert the prior technology-push models by emphasizing that the main innovation driver is, in fact, an unsatisfied market need from which the innovation linearly flows back to research and development (R&D). In these models, the primary role of R&D is hence to meet market demand. Considering the linear models of technology-push and market-pull too simplistic, third generation or *coupling models* (Kline and Rosenberg, 1986) conceptualize innovation as an interactive and not necessarily continuous process involving loops and intra- and inter-organizational linkages, which eventually matches unfilled market needs and technology. *Integrated or parallel models* of the fourth generation involve overlapping activities in different functional departments of a firm and their integration with activities of suppliers and customers (Graves, 1987). Fifth generation models of *systems integration and networking* (Rothwell, 1994) extend fourth generation models by emphasizing that the learning takes place within networks of suppliers, collaborators, alliance and joint-venture partners, customers, lead users (von Hippel, 1986), and even competitors.

Implicit in all models mentioned so far is the assumption that the innovation process is *closed*, in the sense that it takes place within the boundaries of a single firm and/or that it is based on proprietary technology and intellectual assets. The *open innovation* paradigm or model (Chesbrough, 2003; West and Bogers, 2014) changes this assumption by explicitly embracing insourcing and/or outsourcing of technology and intellectual assets, spin-offs, and both internal and external paths to the market. Finally, the *triple helix model* (Etzkowitz, 2008) takes the stakeholder view to propose that innovation occurs through complex and repeating interactions among academia (universities), industry (enterprises), and government.

In sum, the driver of the innovation process in general can be internal (available resources including technology and knowledge), external (a market opportunity), or a coupling (match) of internal with external drivers. As for the loci, innovation can occur within a single firm or a network, which may include some or all collaborating firms, suppliers, partners, and customers. We conceptualize open and triple helix innovation as special cases of innovation taking place in networks. The former is characterized by insourcing and outsourcing of intellectual property, spin-offs and/or external ways to the market, the latter by repeating interactions among universities, enterprises, and governments.

Like innovation in general, big data innovation is a two-dimensional phenomenon that includes a process (Coussement et al., 2017; Del Vecchio et al., 2018; Kayser et al., 2018; Bresciani et al., 2021) and the outcome of the process that may take any form, be it a product (Langevin, 2019), service (Troilo et al., 2017; Gao et al., 2020), business process (Liu et al., 2019) or business model

(Seggie et al., 2017; Sorescu, 2017; Chen, 2019; Trabucchi and Buganza, 2020). Analogously, the process of big data innovation is characterized by the sub-dimensions driver (or stimulus) and locus. However, drivers of big data innovation are idiosyncratic (Table 1), given that we distinguish between big data in the narrower (only datasets of unprecedentedly large scope and volume) and the broader sense (which additionally encapsulates big data analytics).

Table 1: Drivers of big data innovation

Driver	Description
Resource push	Big data in its narrower sense is a “digital raw material” that stimulates innovation by its mere presence and potential to generate value (“data push innovation”; Trabucchi and Buganza, 2020).
Big data pull	The need to make big data in its narrower sense actionable stimulates innovation in technologies and tools for descriptive, diagnostic, predictive, and/or prescriptive analytics (Fleckenstein and Fellows, 2018).
Technology push	Invention (i.e., the first-ever occurrence of an idea for a product, process, or solution; OECD and Eurostat, 2018) in big data analytics or technology (i.e., big data in its broader sense) stimulates the innovation process to satisfy some previously latent market needs (cf. Saheb and Saheb, 2020).
Market pull	Recognized but unsatisfied market needs stimulate purposeful innovation in big data technology and analytics (cf. Wright et al., 2019), i.e., big data innovation is “pulled” by the market.
Coupling	Big data innovation is driven by the attempt to match market needs on one side and big data and analytics on the other (“affordances”; De Luca et al., 2021).

Loci of big data innovation are robust to the different meanings of the term big data, so they mirror the loci of innovation in general (Table 2).

Note that neither drivers nor loci of big data innovation presume any specific model of the innovation process since drivers are generally and loci largely applicable irrespectively thereof. For example, the “closed/ambiguous” locus fits any of the first three generation models and “network” any of the subsequent ones.

3 Method

This paper employs a systematic literature review as a transparent, elaborate, and replicable method for the identification, evaluation, and synthesis of extant research (Tranfield et al., 2003; Rousseau et al., 2008; Hodgkinson and Ford, 2014; Fink, 2013). In line with best practices (ibid), we perform our review in the following stages.

3.1 Planning and Execution

Rather than to a single method, systematic reviews refer to a family of methods that aim at drawing holistic conclusions from the literature being reviewed (Boaz et al., 2006). In line with our research question, we aimed at extracting prior conceptual contributions regardless of whether the respective publication is conceptual, empirical, or a literature review. We exclusively aimed at business/management literature and therefore chose ProQuest's ABI/Inform Collection as the academic database to search in, for it is the most comprehensive database of business literature

Table 2: Loci of big data innovation

Locus	Description
Closed/Ambiguous	Big data innovation takes place within a single firm and/or with its proprietary knowledge/technology; alternatively, the locus is ambiguous in the sense that more than one locus may apply (cf. Johnson et al., 2017).
Network	Big data innovation takes place in a network of collaborators, partners, suppliers, and factual or potential customers; network actors share either or both big data in the narrower sense and big data analytics/technology (cf. De Luca et al., 2021).
Open	A special case of big data innovation in networks based on financial transactions for insourcing and/or outsourcing of big data and analytics/technology (cf. Fortunato et al., 2017; Del Vecchio et al., 2018).
Triple Helix	A special case of big data innovation in networks characterized by repeating interactions among universities, industry, and government (cf. Kim and Lee, 2016).

(Gusenbauer and Haddaway, 2020; Gusenbauer, 2022). Given that titles and abstracts are good proxies of a publication's content (Abrahamson and Eisenman, 2008), we searched for publications that contain either the keyword "big data" or "innovation*" in the title and the other keyword either in the title or in the abstract. We searched only for peer-reviewed articles in English published in scholarly journals. No restrictions regarding publication dates were set. Further, we decided to exclude the following items:

- Editorials, opinion papers, and transcripts of interviews or panel sessions.
- Papers in which the keywords are used in a context different from ours, e.g., in which big data analytics such as data mining is used to answer a research question related to innovation in general but not to big data innovation.
- Papers on big data innovation without findings pertinent to its drivers and loci.
- Papers from outlets known for charging publication fees unless listed in the Directory of Open Access Journals (<https://doaj.org>).

Eventually, the search query in ABI/Inform was executed on November 17, 2021, and retrieved a raw sample of 138 items.

3.2 Quality Assurance

We removed one duplicate and read abstracts and partly full texts of the remaining 137 articles to eliminate 68 false positives, i.e., articles that meet the exclusion criteria. Each researcher performed this task independently. Subsequently, we discussed divergent opinions until a unanimous decision had been made. One of the remaining articles was retracted while our paper was in the publishing process and was hence retroactively eliminated. Eventually, we arrived at a final sample of 68 articles.

The final sample was cross-checked with Scopus, where we did not find 33 of 68 articles in the sample, which reinforces our decision to use ABI/Inform.

3.3 Analysis

Given that we aim at a conceptual consolidation and that the collected data (article full texts and extracted main findings) are qualitative, we applied the qualitative analysis techniques of pattern matching and explanation building (Yin, 2014). Pattern matching compares empirically based patterns with the patterns predicted before the data collection (ibid). In our case, it was used to verify and eventually confirm the drivers and loci of big data innovation identified initially (Tables 1 and 2). Subsequently, the final sample was categorized two-dimensionally (i.e., each article was assigned to one driver and one locus) according to the respective main findings (Table 3). Each researcher performed the categorization independently, divergences were discussed, and a unanimous common decision reached. The categorization and article metadata served later as the base for performing our descriptive analysis.

In most cases, the categorization proved straightforward; in ambiguous cases, additional literature was consulted. For example, articles that mention or allude to dynamic capabilities were categorized as “coupling” since dynamic capabilities refer to the building and recombining of knowledge, technology, and resources to sense and seize opportunities arising from a changing environment (Teece et al., 1997). Similarly, a process of big data innovation resembling the technology-organization-environment (TOE) model was assigned to the locus category “network” (unless special cases of open or triple helix innovation were recognized), since the TOE model refers to technologies internal and external to the focal firm (Tornatzky and Fleischer, 1990).

Explanation building is narrative and iterative. It starts from an initial explanation that is repeatedly revised until it eventually matches all data collected (Yin, 2014). In our case, it was used to explain the main findings on each driver of big data innovation separately, along the article categories built, while taking account of loci and theoretical perspectives used.

Table 3: Two-dimensional categorization of the final sample over drivers and loci

	Closed/Ambiguous	Network	Open	Triple Helix
Resource Push		Sorescu (2017); van den Broek and van Veenstra (2018); von Grafenstein et al. (2019)	Mann (2018); Trabucchi and Buganza (2020); Trabucchi et al. (2018)	Mastrangelo (2018); Nayar and Nayar (2014); Xu et al. (2020)
Big Data Pull	Kozak et al. (2021); Liu and Liu (2019); Yamamoto and Lloyd (2019)	Masih and Joshi (2021); Nuccio and Guerzoni (2019)	Ajimoko (2018)	
Technology Push	Akpan et al. (2021); Bharati and Chaudhury (2019); Chai and Shih (2017); Chen (2018); Troilo et al. (2017)	Akhtar et al. (2019); Chen (2019)	Gao et al. (2020)	Frizzo-Barker and Chow-White (2014); Iorio et al. (2020); Lin et al. (2020); Saheb and Saheb (2020); Zhang et al. (2019)

	Closed/Ambiguous	Network	Open	Triple Helix
Market Pull	Seggie et al. (2017)	Coussement et al. (2017); Kayser et al. (2018); Pricop and Orzan (2018); Rodriguez and Da Cunha (2018); Somoza Sanchez et al. (2018)	Fortunato et al. (2017); Wright et al. (2019)	Dubé et al. (2018); Liu et al. (2019); McGregor (2018); Sollins (2019); Sonka (2016)
Coupling	Arthur and Owen (2019); Blackburn et al. (2017); Cather (2018); Dicuonzo et al. (2019); El-Kassar and Singh (2019); Lim et al. (2020); Johnson et al. (2017); Langevin (2019); Mikalef et al. (2019)	De Luca et al. (2021); Lambrou (2016); Lee (2018); Liu et al. (2020); Rejeb et al. (2020); Saldžiūnas and Skyrius (2017); Sati (2017); Yablonsky (2019)	Duncan et al. (2021); Gradeck et al. (2019); Prescott (2016)	Aboelmaged and Mouakket (2020); Bresciani et al. (2021); Chae (2019); Chandy et al. (2017); Del Vecchio et al. (2018); Huang et al. (2018); Kim and Lee (2016)

3.4 Synthesis

Finally, we synthesized our findings within a framework that we have derived from OECD and Eurostat (2018) and Crossan and Apaydin (2010). This framework considers the environment of big data innovation, its determinants at the individual, organizational and processual levels, as well as its both dimensions process (including subdimensions driver and locus) and outcome. We employed the explanatory synthesis method since it is particularly useful in fragmented and methodologically diverse fields such as business/management and because it stresses the role of theory and makes it explicit (Rousseau et al., 2008).

4 Descriptive Analysis

The oldest two articles in the final sample were published in 2014. The number of publications rapidly increased from 2016 to 2019, the year with the highest publication frequency (19 articles; Figure 1). Our database query was executed in November 2021 and consequently cannot include articles published later that year. Nevertheless, the sample distribution suggests a decreasing publication activity since 2019. A potential explanation for this may be a temporal shift of research priorities due to the Covid-19 pandemic or a gradual saturation of the research topic.

The articles in the final sample have been published by a total of 175 authors, multiple occurrences included. Each author's country of affiliation at the time of publication was counted once, irrespective of whether first, second, or corresponding author. If an author had multiple affiliations, each of them was counted. The largest share of authors is affiliated in Italy (17%), followed by the US (14%) and China (13%; Figure 2).

This sharply contradicts the geographical distribution of big data patent applications. According to Saheb and Saheb (2020), 78% of patent applications is filled in China and 8% in the US,

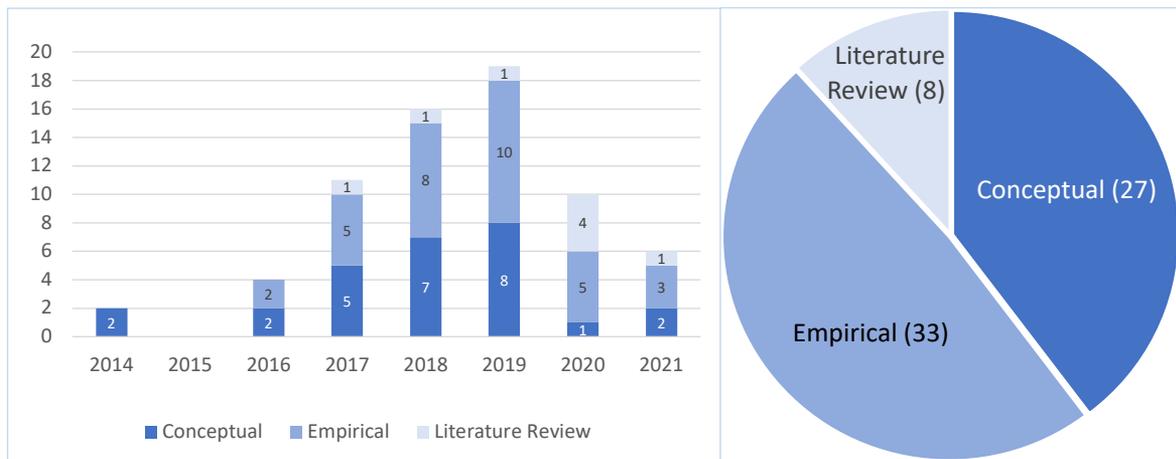


Figure 1. Distribution of the final sample over years and article categories

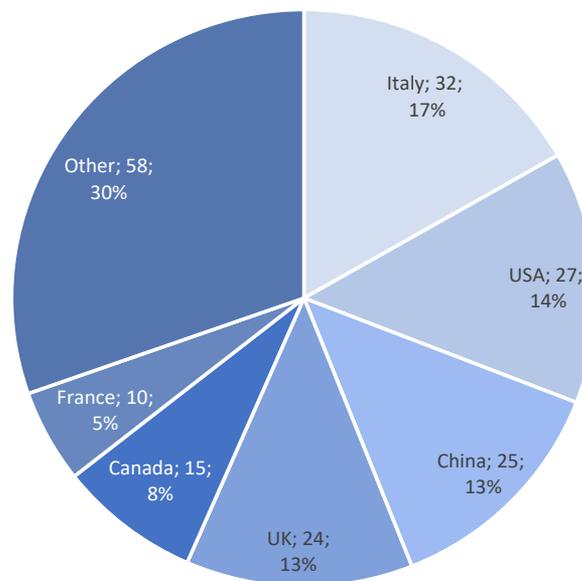


Figure 2. Distribution of authors' country of affiliation in the final sample

whereas no big data patent applications were filled in Italy. Different propensity to collaboratively publish may partly explain why authors affiliated in Italy are comparatively overrepresented. For articles authored by at least one person affiliated in Italy, the average number of authors with Italian affiliations is 3.2 compared to 1.6 in the case of the US. However, this ratio is 2.5 in the case of China, so it fails to explain why authors affiliated in China are heavily underrepresented. The publication language is more likely an explanation here since most big data researchers affiliated in China presumably prefer to publish in Mandarin rather than in English. In line with this, our brief search with the same keywords in the Chinese Social Sciences Citation Index (CSSCI; <http://cssci.nju.edu.cn>) retrieved a raw sample of 183 publications on big data innovation in Mandarin.

The distribution of the final sample over categories of drivers and loci is presented in Figure 3. This distribution identifies coupling as the *single* driver of big data innovation most frequently occurring in extant literature (27 articles), with market pull and technology push being jointly

second (13 each). Resource push and big data pull are elaborated on less frequently (9 and 6 articles, respectively).

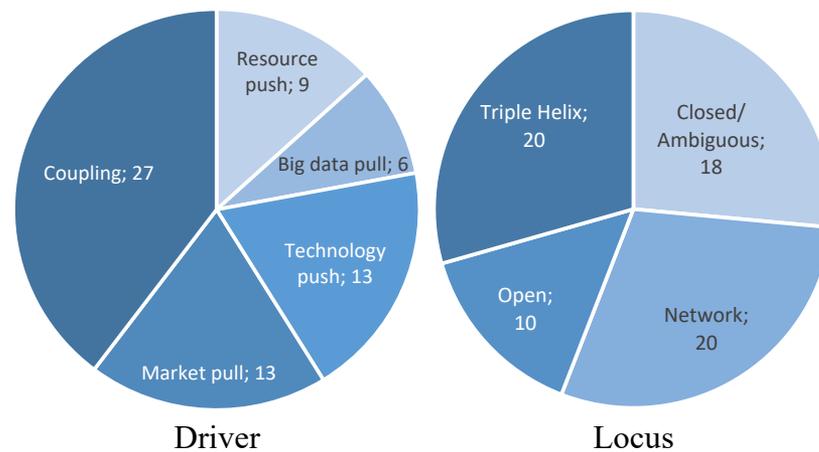


Figure 3. Distribution of the final sample over drivers and loci

One thing should be noted here, however. Resource push, big data pull, and technology push are *endogenous* drivers in the sense that they all imply the stimulus to innovate residing internally in the mere existence of big data (resource push), which pulls technological innovation in analytics (big data pull) so as to be commonly pushed to the market (technology push). All endogenous drivers combined account for the biggest share of literature (28 articles) in the final sample. Market pull (13 articles) represents *exogeneous* innovation stimulus residing in unsatisfied market needs, whereas coupling (27 articles) represents the dynamic interplay of exogenous and endogenous stimuli. In other words, the less mere data and analytics and the more market need are involved, the fewer articles are there in the final sample.

As for loci, a closed/ambiguous innovation process is rather an exception (18 articles) since most articles depict big data innovation as a cooperative process that takes place either in networks (20 articles), in the form of open innovation (10 articles), or in triple helix constellations (20 articles).

5 Drivers of Big Data Innovation

5.1 Resource Push

In the simplest scenario, big data serves as a digital raw material that stimulates innovation by its mere presence and (assumed) potential to generate value through “data push innovation” (Trabucchi and Buganza, 2020). From extant literature, sharing in innovation networks appears to be the dominant option for resource push, in which two or more organizations combine their big data to innovate (van den Broek and van Veenstra, 2018; von Grafenstein et al., 2019), particularly so new or improved business models (Sorescu, 2017). As a special case of resource push in networks, sharing occurs according to the triple helix model (Nayar and Nayar, 2014; Mastrangelo, 2018; Xu et al., 2020). For the innovative capacity of actors to increase through sharing, however, unconnected islands of big data need to be connected (Mastrangelo, 2018), for which “small data” that provide deeper, contextual insights may be particularly beneficial (Xu et al., 2020). The hypothesis that big data is a non-rivalrous good as proposed by OECD (2015) and echoed in von Grafenstein et al. (2019) – so that sharing provides value to the receiver

without diminishing the value for the provider – is partly objected to. Taking a political economy perspective, Mann (2018) claims that many development projects in Africa extract data for expert analysis in developed countries, which ultimately deprives African countries from a source of knowledge, revenue, and power.

Alternatively, an organization may sell big data to others (Trabucchi et al., 2018; Trabucchi and Buganza, 2020) rather than share it. This implies open innovation (Chesbrough, 2003) and should therefore be understood more generally, as trading big data, which may include either or both outbound open innovation (outsourcing/selling) or inbound open innovation (insourcing/buying; West and Bogers, 2014).

In sum, extant literature implies two generic strategies for big data innovation driven by resource push, sharing and trading. Using big data within a single organization to innovate on its own as a potential third generic strategy is not explicitly depicted yet not explicitly denied either. These three generic strategies are not mutually exclusive and could hence be mixed. While sharing “raw” big data is the most frequently depicted strategic option, it simultaneously raises issues of incentives and governance. Particularly in the EU, there are more obstacles to than incentives for sharing big data (von Grafenstein et al., 2019), which forces firms to balance potential gains and compliance with data protection legislation. Big data governance is hence highly dependent on regulatory pressures (van den Broek and van Veenstra, 2018), which clearly frames big data innovation with institutional theory (Scott, 2014) and isomorphic pressures to comply with the rules and beliefs prevailing in the environment. However, beyond institutional theory and irrespectively of the generic strategy pursued, theoretical foundations of resource-pushed big data innovation have been barely elaborated on so far.

5.2 Big Data Pull

Big data provides opportunities for organizations to learn and take well-informed (or at least better-informed) actions, but it is inactive and has no value in and of itself (Yamamoto and Lloyd, 2019). The need to make big data actionable through context and presentation pulls – i.e., drives or stimulates – innovation in technologies and tools for accessing, collecting, and managing big data for descriptive (“What happened?”), diagnostic (“Why did it happen?”), predictive (“What will happen?”), and prescriptive analytics (“How can we make it happen?”; Fleckenstein and Fellows, 2018). Consequently, extant literature particularly stresses the need for new or improved analytics for marketing, digital platforms, and Internet of Things (IoT) services.

Marketing already heavily relies on digital infrastructure to increase the speed, scale, adaptability, and precision of content management, advertising, and customer relationship management (Grimpe et al., 2017). Further innovation in big data analytics is required to provide just-in-time information feeding of agile marketing (constant validation of ideas about new features, products, or services with customer feedback) on an ongoing, large scale (“big testing”; Yamamoto and Lloyd, 2019). In particular, analysis of big data from social media may help manufacturers of health/organic foods to increase sales through constant improvement of their products according to customers' expectations, preferences, sentiments, and perceptions (Masih and Joshi, 2021). Additionally, big data-based swarm intelligence machine learning may be applied to churn management, i.e., managerial decision-making regarding customers who are considering leaving a given company's customer base (Kozak et al., 2021). Innovation in analytics applied to digital platforms may assist companies in identifying and seizing opportunities that latently exist in big data and hence compete on the ability to successfully leverage it. For example, dating sites should have sophisticated matching mechanisms, search engines should correctly predict consumers' interests, and e-commerce portals should effectively suggest items of complementary interest to buyers

(Nuccio and Guerzoni, 2019). Finally, for IoT environments, business intelligence applications are required to merge and analyze big data from various sources for improved process performance and decision-making regarding smart services such as banking, governance, power generation, city lighting, health care, and mobility (Liu and Liu, 2019).

According to one half of the articles in the subsample, the development of new or improved technologies and analytical tools takes place cooperatively, in networks (Nuccio and Guerzoni, 2019; Masih and Joshi, 2021), or in the form of open innovation (cloud-based services for big data analytics; Ajimoko, 2018). The other half of articles on big data pull do not explicitly advocate cooperation, yet they at least employ publicly available customer data (Kozak et al., 2021), IoT (Liu and Liu, 2019), and social media (Yamamoto and Lloyd, 2019) to answer the respective research question.

The diffusion of innovations theory (Rogers, 2003) is the major theoretical frame for big data pull. In this context, extant literature stresses the role of determinants of innovation diffusion such as executives' support (Ajimoko, 2018), positive and active enterprise culture (Liu and Liu, 2019), or dynamic capabilities (Kozak et al. 2017). In addition, several sources in this subsample refer to the resource-based theory of the firm (Barney, 1991) by asserting or implying that new or improved big data technologies/tools may provide a source of competitive advantage for the firms which use them (Ajimoko, 2018; Nuccio and Guerzoni, 2019; Yamamoto and Lloyd, 2019; Kozak et al., 2021). However, this is open to debate given that a company can gain a sustainable advantage over its rivals only by having proprietary technology that they cannot have. In contrast, big data technology and analytics are likely to diffuse until they eventually become an infrastructural technology in the sense of Carr (2003).

5.3 Technology Push

Technology push refers to the innovation process triggered by the invention in big data analytics or technology, which is eventually converted into new products or services to satisfy some previously latent market needs. Given that invention can legally be protected through patents, dominant themes in this subsample are big data patent output, potential areas of the patent application, and factors supportive of their assimilation.

Most big data patents pertain to information retrieval, database structures, and file system structures within the fields of computer science and medicine (Saheb and Saheb, 2020). A vast majority of 78% of patent applications is filled in China. The second largest number of applications is filled in the US (8%) and the third largest in South Korea (7%), with all other jurisdictions contributing less than 0.5% each. The links between patent applications and academic work are weak since only approx. 14% of patent applications in the US and less than 1% in China cite scientific articles (ibid). No academic work subject to our review has attempted to explain these spatial patterns in big data patent activity. Generally, however, Lin et al. (2020) frame big data patent activity by upper echelons theory (Hambrick and Mason, 1984) and social network theory (Freeman, 1978) to prove that patent output is determined by managerial power and the centrality of a given firm within research and innovation networks. In addition, Zhang et al. (2019) reveal an invertedly U-shaped relationship between patent output and the path length from a firm's knowledge to the knowledge of its partners in an innovation network.

The articles on big data patent applications either conceptualize potential use cases or report first use cases worthwhile mimicking. These cases include the transformation of business models (Chen, 2019), genomic sequencing (Frizzo-Barker and Chow-White, 2014), resource usage in medical institutions (Chen, 2018), steering of business activities during lockdowns caused by Covid-19 and potentially future pandemics (Akpan et al., 2021), and replacement of traditional

by big data-based statistics in the tourism industry (Iorio et al., 2020). Methods of big data analytics seem to have already found their way into academic research as well, which is implied by their application in several articles in our final sample (e.g., Zhang et al., 2019; Saheb and Saheb, 2020) as well as in 15 of false positives that apply big data analytic to answer research questions unrelated to ours. However, in the context of scientific research, Chai and Shih (2017) warn of too high a reliance on big data analytics and urge to use it as a supplement to existing methods, rather than as a substitute.

Successful assimilation of big data patents is contingent upon legal regulations (Frizzo-Barker and Chow-White, 2014), organizational culture, structures, processes, roles, and capabilities (Troilo et al., 2017), and social capital of a firm, particularly trusted and reliable relationships with business partners (Bharati and Chaudhury, 2019). Within alliances and joint ventures, the assimilation is positively affected by the personal power of key actors (expert power and referent power) but negatively by their position power (coercive and manipulative; Akhtar et al., 2019). In addition, new, big data-related professions are required, such as data scientist and data analysts (Iorio et al., 2020). Ultimately, Gao et al. (2020) propose that hosting big data applications in containerized clouds would reduce the complexity of deployment and operation and thus accelerate diffusion and assimilation.

In sum, extant research on technologically pushed big data innovation focuses either on the trigger of the process (i.e., patents) or its end (areas of potential application and facilitators of successful assimilation). Integrative frameworks are missing, equally so empirical studies about the core of the process, i.e., the development of big data innovations that truly meet some latent market needs.

5.4 Market Pull

Market pull as a driver corresponds to a purposeful application of big data and analytics to innovate in response to market opportunities arising from known but unsatisfied needs of enterprises and/or society (Wright et al., 2019). The literature on this driver is essentially bisected in empirical studies, which elaborate on big data innovation successfully catering to recognized market needs, and conceptual papers, which identify market needs yet to be catered to.

Empirical studies focus on making use of big data from social media to satisfy general social needs such as precision public health (McGregor, 2018), affordable foods for the poor (Dubé et al., 2018), responsible innovation of digital platforms (Liu et al., 2019), or the needs of enterprises to generate (additional) value. Specifically, data from social media may be employed to design TV shows (Fortunato et al., 2017), identify lead users, and generate ideas for innovation in the entertainment industry by analyzing the whole population rather than samples of users (Somoza Sanchez et al., 2018), or to incorporate only constructive consumers' feedback in the innovation process (Coussement et al., 2017).

Conceptual papers assert that market opportunities for big data-driven innovation to generate business value abound (Kayser et al., 2018), particularly so in agriculture (Sonka, 2016), sustainable supply chain innovation (Rodriguez and Da Cunha, 2018), mitigation and containment of epidemics (Sollins, 2019), or (B2B) marketing (Pricop and Orzan, 2018; Wright et al., 2019). Prominent channels to generate business value are the rapid verification of managerial hypotheses and adaptation of business models (Seggie et al., 2017), optimization of processes and resource deployment (Rodriguez and Da Cunha, 2018; Sollins, 2019), and/or improved (real-time) decision making (Sonka, 2016; Pricop and Orzan, 2018; Wright et al., 2019).

Major challenges result from security and privacy requirements (Sollins, 2019) and the mere characteristics of big data (volume, velocity, variety, valence, veracity, variability, and vagueness),

both of which impede the integration of big data into the organizational culture, flows and processes (Pricop and Orzan, 2018). In addition, knowledge, resources, and capabilities needed to innovate are scattered across many actors since the literature on market pull sees the process of big data innovation to take place cooperatively, be it in networks (Coussement et al., 2017; Kayser et al., 2018; Pricop and Orzan, 2018; Rodriguez and Da Cunha, 2018; Somoza Sanchez et al., 2018), as open innovation (Fortunato et al., 2017; Wright et al., 2019) or in triple helix ecosystems (Sonka, 2016; Dubé et al., 2018; McGregor, 2018; Liu et al., 2019; Sollins, 2019).

Especially Sonka (2016), Coussement et al. (2017), Rodriguez and Da Cunha (2018), Somoza Sanchez et al. (2018), and Wright et al. (2019) identify the lead user theory (von Hippel, 1986; Franke et al., 2006), the resource-based theory of the firm (Barney, 1991), absorptive capacity (Cohen and Levinthal, 1990), and dynamic capabilities (Teece et al., 1997) as major theoretical foundations of market-pull driven big-data innovation. Additionally, Kayser et al. (2018) conceptually derive a phased, systematic innovation process to retrieve value from big data innovation by satisfying existing market needs. However, empirical studies on successful cases of big data innovation pulled by market needs tend to be phenomenological rather than ontological. For big data innovation to tap into opportunities arising from unsatisfied market needs, future research should hence focus on theoretical generalization and operationalization of practical insights.

5.5 Coupling

Coupling as a driver refers to attempts to match market needs on the one side and big data and analytics on the other. Opportunities for coupling are arising in diverse areas such as solving social and environmental problems of developing countries (Chandy et al., 2017), the development of policies to balance economic benefits from casinos and control of gambling addiction (Kim and Lee, 2016), the integration of global supply chains (Sati, 2017), pricing models of insurers (Cather, 2018), risk management of small banks (Dicuonzo et al., 2019), the commercialization of microfinance (Langevin, 2019), decision-supporting systems for agriculture (Duncan et al., 2021), and digital twins (Lim et al., 2020). Purposeful building and recombining of resources and capabilities to sense and seize opportunities in a changing organizational environment are generally captured by the concept of dynamic capabilities (Teece et al., 1997), which is consequently the most prominent theme in the literature on coupling as a driver of big data innovation. Regarding the sensing of opportunities, extant literature suggests spotting patterns of customer behavior (De Luca et al., 2021), deriving insights from customer service (Rejeb et al., 2020), and exchanging rewards for customers' consent to disclose commercially valuable information (Arthur and Owen, 2019).

Seizing opportunities challenges technology and organization (esp. decision making and organizational culture; Saldžiūnas and Skyrius, 2017) and requires ethical and regulatory responsive data governance (Arthur and Owen, 2019). Developing big data innovation capabilities equally requires adaptive knowledge management (Lambrou, 2016) and real-time, data-driven market ambidexterity (De Luca et al., 2021), i.e., appropriate management of the exploitation-exploration tension commonly acknowledged in innovation management (Andriopoulos and Lewis, 2009). To explore emerging opportunities, firms need to embrace the volume, variety, and velocity (3Vs) of big data (Johnson et al., 2017), which will disrupt R&D and innovation management (Blackburn et al., 2017) through big data technology assimilation (Arthur and Owen, 2019; Aboelmaged and Mouakket, 2020) and the ultimate shift towards machine led and machine governed innovation process (Yablonsky, 2019).

Overcoming challenges of coupling is asserted to enable firms to gain a competitive advantage (Prescott, 2016; El-Kassar and Singh, 2019) through improved performance regarding both radical and incremental big data innovation (Mikalef et al., 2019). However, suggestions on how to overcome challenges essentially boil down to a phased big data technology delivery system (Huang et al., 2018) and a big data-based innovation cycle (Lee, 2018) with simultaneous resorting to sources of data, knowledge, and technology external to the focal firm. Consequently, a vast majority of the extant literature on coupling sees the process of big data innovation to take place in the form of co-innovation (Bresciani et al., 2021), open innovation (Prescott, 2016; Del Vecchio et al., 2018; Duncan et al., 2021) or triple helix innovation (Kim and Lee, 2016; Chandy et al., 2017; Huang et al., 2018; Aboelmaged and Mouakket, 2020; Bresciani et al., 2021), within big data innovation ecosystems (Chae, 2019) or innovation networks in general (Lambrou, 2016; Saldžiūnas and Skyrus, 2017; Sati, 2017; Lee, 2018; Liu et al., 2020; Rejeb et al., 2020; De Luca et al., 2021).

In sum, challenges of coupling are abundant, yet opportunities are as well, so the expected gains sound promising. Theoretical fundamentals of this stream of literature are dynamic capabilities, organizational ambidexterity, and open innovation, yet practical suggestions remain rare and generic rather than capable of being acted on. Not surprisingly, surveys reveal that particularly SMEs do not effectively use big data to innovate since two thirds of them do not even use customer relationship management systems (Liu et al., 2020). In accordance with this, a majority of surveyed 5,700 practitioners in manufacturing consider big data to have neither high importance for innovation nor for their business activities (Gradeck et al., 2019).

6 Synthesis

This section employs the framework visualized in Figure 4 to synthesize our findings. Like innovation in general (OECD and Eurostat, 2018), big data innovation is a two-dimensional phenomenon that includes big data innovation as a process and big data innovation as an outcome. The process necessarily precedes the outcome and contains sub-dimensions of driver and locus (the inner right portion of Figure 4). We adopt the suggestion by Crossan and Apaydin (2010), who group the determinants of the innovation process in general at the individual (or leadership), organizational, and business process levels and group the determinants of big data innovation analogously (the inner left portion of Figure 4). All determinants and both dimensions of big data innovation are placed within the environment (the outer portion of Figure 4).

In line with institutional theory (Scott, 2014), the **environment** exerts isomorphic pressures on big data innovation to align data governance with ethical and legal requirements. While these pressures are common irrespectively of the driver, they particularly hamper open innovation and innovation in networks on the one hand yet promote big data innovation in triple helix constellations on the other (Frizzo-Barker and Chow-White, 2014; van den Broek and van Veenstra, 2018; Arthur and Owen, 2019; Sollins, 2019; von Grafenstein et al., 2019).

At the **individual or leadership level**, upper echelon theory (Hambrick and Mason, 1984) connects individual characteristics and behavior of executives with organizational outcomes. In context of big data innovation, leadership and managerial power determine big data patent output (Lin et al., 2020), assimilation of big data technology (Akhtar et al., 2019) and, ultimately, the diffusion of the innovation outcome (Ajimoko, 2018). In turn, taking advantage of big data innovation requires hands-on, agile leadership that sifts through opportunities and matches them with aggregated resources and talent (Nayar and Nayar, 2014).

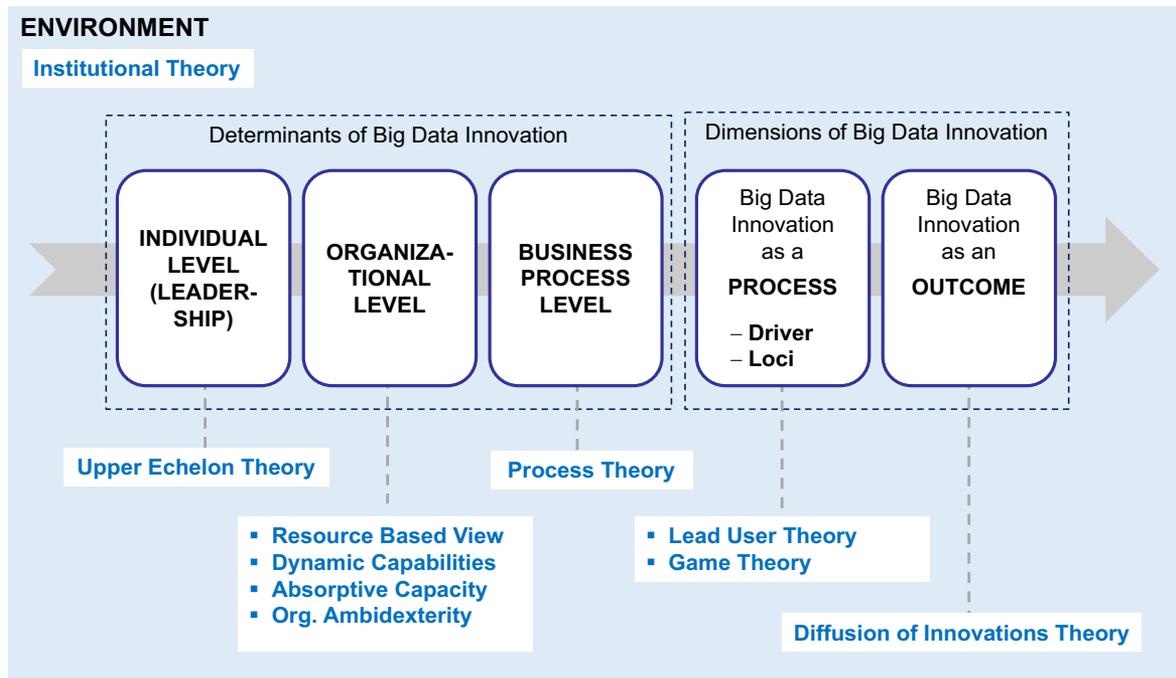


Figure 4. Theoretical perspectives towards big data innovation in extant literature

Determinants of big data innovation at the *organizational level* are comparatively well-researched and theoretically well-grounded. In order to sense and seize opportunities in a changing environment, purposeful building and recombining of resources and capabilities (dynamic capabilities; Teece et al. 1997) is as much required as appropriate management of the exploitation-exploration tension (organizational ambidexterity; Andriopoulos and Lewis, 2009) and the ability to identify, assimilate, transform, and apply external knowledge typically scattered over innovation networks (absorptive capacity; Cohen and Levinthal, 1990). In addition, several sources (Prescott, 2016; Sonka, 2016; Ajimoko, 2018; Rodriguez and Da Cunha, 2018; El-Kassar and Singh, 2019; Nuccio and Guerzoni, 2019; Wright et al., 2019; Yamamoto and Lloyd, 2019; Kozak et al., 2021) refer to the resource-based theory of the firm (Barney, 1991) and suggest that big data innovation may provide a source of sustainable competitive advantage.

Organizational process theory (Van de Ven and Poole, 1995) is the major theoretical frame for big data innovation at the *business process level*. Generally, process theory explains how an organization transforms inputs into outputs and how it changes and develops. In this context, big data innovation both challenges and improves decision-making and resource allocation (Sonka, 2016; Saldžiūnas and Skyrius, 2017; Seggie et al., 2017; Troilo et al., 2017; Pricop and Orzan, 2018; Rodriguez and Da Cunha, 2018; Sollins, 2019; Wright et al., 2019; Lim et al., 2020; Duncan et al., 2021).

In terms of *big data innovation as a process*, particularly the literature on market-pull and coupling heavily draws on the lead user theory (von Hippel, 1986; Franke et al., 2006) since lead user groups, user-generated big data, and social media serve as the major source of ideas for new or improved products, services, and business models (Sonka, 2016; Coussement et al., 2017; Fortunato et al., 2017; Dubé et al., 2018; McGregor, 2018; Rodriguez and Da Cunha, 2018; Somoza Sanchez et al., 2018; Trabucchi et al., 2018; Liu et al., 2019; Wright et al., 2019; Masih and Joshi, 2021). References to game theory (McCain, 2010) are not explicit yet conspicuous

given our findings which indicate that cooperation is the dominant interactive strategy for big data innovation.

At this level of analysis, our review provides clear evidence that big data simultaneously drives the innovation process in multiple ways: as abundant digital raw material with the potential to generate value (resource push), as a market opportunity for big data technologies and analytical tools (big data pull), as a subject to the invention which may eventually be converted into products or services to satisfy latent market needs (technology push), as a stimulus to innovate in response to existing but unsatisfied market needs (market pull), and as a means to match existing market needs with technologies and analytical tools (coupling). The extent to which these drivers are represented in reviewed literature is different, however. Big data innovation exogenously driven by market needs (market pull) is far less frequently elaborated on than big data innovation driven endogenously (resource push, big data pull, and technology push) or dually, by the interplay of exogenous and endogenous factors (coupling; cf. Figure 3).

The process of big data innovation primarily takes place cooperatively, as open innovation, triple helix innovation, or in idiosyncratic networks of actors, including some or all partners, suppliers, users, and competitors. There are several potential explanations for this observation. First, linkages and related recombination of resources, technology, and knowledge that exist beyond a single organization's boundaries play a fundamental role in the innovation process in general (Savino et al., 2017), which is likely only stressed when it comes to big data innovation in particular. Second, the hypothesis that big data is a non-rivalrous good (OECD, 2015) may hold, so that sharing mere big data (resource push) is a Pareto improvement in which some agents may gain, but no agent will lose. Third, sharing may be a result of competitive failure (Campbell-Hunt, 2000; Markides and Charitou, 2004), in which cooperation is a suboptimal strategy since no competitor can gain competitive or at least comparative advantage yet given that cooperation is followed by virtually all competitors, no competitor can end up having a disadvantage either. Finally, the motive for cooperation with governments (especially but not exclusively in triple helix networks) may particularly be the pressure to align big data governance with legal requirements, or to influence them in the first place.

The diffusion of innovations theory (Rogers, 2003) explains how, why, and at what rate **big data innovation as an outcome** spread over the categories of adopters (innovators, early adopters, early majority, late majority, and laggards; *ibid*). On one hand, the diffusion of big data innovations is accelerated by executives' support (Ajimoko, 2018; Akhtar et al., 2019), positive and active enterprise culture (Troilo et al., 2017; Liu and Liu, 2019), and cloud-based big data applications (Gao et al., 2020). On the other hand, however, the diffusion is slowed down by the mere characteristics of big data ("Vs" such as volume, velocity, and variety; Pricop and Orzan, 2018), security and privacy requirements (Frizzo-Barker and Chow-White, 2014; van den Broek and van Veenstra, 2018; Sollins, 2019; von Grafenstein et al., 2019), and the lack of new, big data-related professions (Iorio et al., 2020). While Trabucchi and Baganza (2020) observe that big data innovation driven by resource push is *comparatively* less diffused, extant literature remains generally inconclusive about the *absolute* spread of big data innovations. However, the skepticism of most practitioners being given (Gradeck et al., 2019), it is safe to assume that big data innovation is presently far away from reaching "the chasm" (Moore, 1999), i.e., the critical mass of early majority.

7 Future Research Directions

The preceding synthesis implies a variety of opportunities for further research on big data innovation, which we present along the elements of Figure 4. At the level of **environment**, legislation on data governance requires increased attention of both researchers and policy makers to balance social gains and risks of big data innovation, particularly so in healthcare (Frizzo-Barker and Chow-White, 2014). So far, innovation policies have primarily provided incentives for basic research and technology absorption within a state or territory, yet big data innovation requires policy tools that go beyond the national context (cf. Bathelt et al., 2017). Our review has identified very few pieces of research supportive for developing such policy tools, except for Sollins (2019) who suggests iterative review and redesign at the intersection of three distinct objectives (social, economic, and technological) underlying big data innovation.

Determinants at the **individual or leadership level** seem to be the least researched area of big data innovation. We see two major themes at this level: leadership regarding data governance and ethics, and leadership regarding the management of the exploration-exploitation tension. A firm's standards of governance and ethics for big data innovation need to be both adaptive and responsive (Arthur and Owen, 2019). However, they also need to go beyond mere legal compliance; not everything that is legal is necessarily considered ethical. As for the exploration-exploitation tension, recent research proves that narcissistic CEOs induce a firm's emphasis on exploitation to the detriment of exploration (Steinberg et al., 2022), while exactly the latter is associated with big data innovation. Whether neglecting big data innovation as a form of exploration affects narcissistic CEOs only is worth elaborating on. The skepticism of most practitioners towards big data innovation (Gradeck et al., 2019) may namely be inherited from the leadership level rather than built intrinsically.

At the **organizational level**, innovation is considered a source of competitive advantage (Tushman and O'Reilly, 1996). Extant literature attributes the same potential to big data innovation, yet – except for Prescott (2016) – provides little evidence to support this claim. Unless proprietary, big data technology and analytics will progressively diffuse and may be adopted by any competitor at a certain point. The advantage gained from them may consequently prove erodible and hence temporal rather than sustainable (cf. Carr, 2003). In addition, further research is needed to prove whether big data innovation enables pioneers (cf. Lieberman and Montgomery, 1988) or late movers (cf. Shankar et al., 1998) to gain competitive advantage, and to elaborate on both mediating and moderating factors that affect the relationship between timing and competitive advantage. A case in point is Facebook, in whose business model big data and analytics represent key resources (cf. Osterwalder and Pigneur, 2010). Facebook was a latecomer to the social media market, but it reportedly outperformed pioneers due to the mastery of public relations rather than due to big data and analytics (Press, 2018).

In addition to challenges to resource allocation and decision-making (Chen et al., 2013; Chen and Zhang, 2014; Troilo et al., 2017; Del Vecchio et al., 2018; Wright et al., 2019), researchers' attention at the **business process level** deserves the solution of challenges and dilemmas posed to management of big data innovation projects. The fact that big data innovation primarily takes place cooperatively exerts qualitative and quantitative pressures on project stakeholder management. Additionally, time-to-market is critical, yet the progress of big data innovation projects typically depends on the unpredictable achievement of learning events (cf. Dougherty et al., 2013).

Regarding **big data innovation as a process**, we observe that empirical studies tend to be phenomenological or micro-theoretical rather than ontological and/or practically oriented.

Consequently, there is a need for both theory-building case studies and best practices generalized and operationalized for being acted on, especially so when it comes to comparatively neglected purposeful innovation in big data technology and analytics to meet existing but unsatisfied market needs (“market pull”). In addition, our review is limited regarding the subdimensions of the innovation process in the sense that we only elaborate on drivers and loci. We thus particularly postpone research questions related to the subdimension “level”, which outlines interfaces and interactions among the individual, group, and firm innovation processes (Crossan and Apaydin, 2010). Big data innovation requires “unleashing” of individual and group level innovation processes (cf. De Smet and Gagnon, 2018), yet extant literature remains mute about how to federate them at the firm level.

Another recognized limitation of our review is that we refrain from elaborating on **big data innovation as an outcome**. However, particularly the magnitude of the outcome (i.e., the degree of newness) deserves further research attention. Whereas radical innovations trigger fundamental changes of organizational practices which (may) result in the transformation of individual firms or whole industries, incremental innovations induce marginal changes that primarily reinforce the existing organizational capabilities (Gopalakrishnan and Damanpour, 1997). Several articles in our final sample, especially conceptual ones, claim that big data innovation may represent a source of competitive advantage, and hence imply a radical magnitude. In contrast, our anecdotal impression is that empirical papers rather depict cases of incremental innovation. This is not necessarily inconsistent, however, given the importance of ambidexterity in managing both radical and incremental innovation (Tushman and O’Reilly, 1996; Andriopoulos and Lewis, 2009).

Finally, our paper is limited to reviewing the literature in English only, whereas we collaterally find indications that there is a comparably large body of literature on big data innovation in Mandarin. A review of the literature in Mandarin presented in English seems therefore a promising project for teams staffed with researchers fluent in both languages.

8 Conclusion

In sum, our literature review makes four major contributions. First, it demonstrates an overall simultaneity and diversity of drivers of big data innovation, yet it concomitantly reveals that big data innovation has been so far primarily driven by attempts to generate value from mere data, to improve analytics, or to commonly push data and technology to the market – but not to meet existing market needs unless they can be matched by existing data and technology. Put bluntly, we reveal that the discourse on big data innovation needs to learn market pull from “traditional” innovation management. Existing market needs for the application of big data and analytics have been neither well understood nor well served so far. In the sense of Sondergaard’s analogy (cited in Yablonsky, 2019), the world accumulates the oil of the 21st century and improves combustion engines, yet cars, trucks, boats, stand-alone power generators and other opportunities for their common application and commercialization are still blurred or neglected.

Second, we observe that big data innovation is primarily a cooperative process. This observation does not contradict the fact that we have assigned several articles to the category “closed/ambiguous” (Table 3). These articles neither advocate nor depict big data innovation as a cooperative process, yet they do not explicitly deny its cooperative character either. They simply rather focus on determinants and challenges of big data innovation within an individual firm. For example, Blackburn et al. (2017), Chai and Shih (2017), Liu and Liu (2019), and Yamamoto and Lloyd (2019) underline the need for R&D departments to take advantage of big data. Other articles in this category describe success stories and potential for big data adoption

from an internal view (Cather, 2018; Arthur and Owen, 2019; Bharati and Chaudhury, 2019) and/or derive normative recommendations (Johnson et al., 2017; Troilo et al., 2017; Mikalef et al., 2019).

Third, we identify and coherently synthesize major theoretical perspectives taken towards big data innovation in extant literature. Although Blackburn et al. (2017, p. 43) have predicted that “big data and big data analytics will have significant implications for R&D and innovation management in the next decade”, our review demonstrates that in fact the opposite is true amid of what the next decade was back then – in the sense that the theoretical lenses of “traditional” innovation management such as dynamic capabilities, absorptive capacity, and organization ambidexterity have framed the research on big data innovation so far. Proprietary theoretical contributions are limited to very few fragmented concepts such as phased innovation process (Kayser et al., 2018), big data technology delivery system (Huang et al., 2018), and big data-based innovation cycle (Lee, 2018). Combined with a lack of focus on existing market needs, this may explain the observation by Zhang et al. (2019) that big data innovation has failed to achieve a big impact anywhere except for the IT industry so far.

Fourth, we identify major directions for future research to build upon our findings, address limitations of our research, expand theory, and replicate our research in another geographic context. Our impression is that the publication activity on big data innovation has increased since this paper went into the publishing process, which supports our assumption that the drop that we observed in 2020-2021 may have been attributable to a temporal shift of research priorities due to the Covid-19 pandemic. In any case, many more exciting studies on big data innovation are warranted and yet to come.

Acknowledgement

The authors received no financial support for the research, authorship, and/or publication of this article.

9 References

- Aboelmaged, M., & Mouakket, S. (2020). Influencing Models and Determinants in Big Data Analytics Research: A bibliometric Analysis. *Information Processing & Management*, 57(4), 1. DOI:10.1016/j.ipm.2020.102234
- Abrahamson, E., & Eisenman, M. (2008). Employee-Management Techniques: Transient Fads or Trending Fashions? *Administrative Science Quarterly*, 53(4), 719–744. DOI: 10.2189/asqu.53.4.719
- Ajimoko, O. (2018). Considerations for the Adoption of Cloud-based Big Data Analytics in Small Business Enterprises. *Electronic Journal of Information Systems Evaluation*, 21(2), 63–79.
- Akhtar, P., Khan, Z., Rao-Nicholson, R., & Zhang, M. (2019). Building Relationship Innovation in Global Collaborative Partnerships: Big Data Analytics and Traditional Organizational Powers. *R&D Management*, 49(1), 7–20. DOI:10.1111/radm.12253
- Akpan, I. J., Soopramanien, D., & Kwak, D.-H. (Austin). (2021). Cutting-edge Technologies for Small Business and Innovation in the Era of COVID-19 Global Health Pandemic. *Journal of Small Business and Entrepreneurship*, 33(6), 607–617. DOI:10.1080/08276331.2020.1799294
- Andriopoulos, C., & Lewis, M. W. (2009). Exploitation-exploration Tensions and Organizational Ambidexterity: Managing Paradoxes of Innovation. *Organization Science*, 20(4), 696–717. DOI:10.1287/orsc.1080.0406

- Arthur, K. N. A., & Owen, R. (2019). A Micro-ethnographic Study of Big Data-Based Innovation in the Financial Services Sector: Governance, Ethics and Organisational Practices. *Journal of Business Ethics*, *160*(2), 363–375. DOI:10.1007/s10551-019-04203-x
- Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2019). Big Data Adoption: State of the Art and Research Challenges. *Information Processing & Management*, *56*(6), 1-18. DOI:10.1016/j.ipm.2019.102095
- Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. DOI: 10.1177/014920639101700108
- Bathelt, H., Cohendet, P., Henn, S., & Simon, L. (2017). Innovation and Knowledge Creation: Challenges to the Field. In H. Bathelt, P. Cohendet, S. Henn, & L. Simon (Eds.), *The Elgar Companion to Innovation and Knowledge Creation* (pp. 1–21). Cheltenham, UK : Edward Elgar Publishing.
- Bharati, P., & Chaudhury, A. (2019). Assimilation of Big Data Innovation: Investigating the Roles of IT, Social Media, and Relational Capital. *Information Systems Frontiers*, *21*(6), 1357–1368. DOI:10.1007/s10796-018-9839-6
- Blackburn, M., Alexander, J., Legan, J. D., & Klabjan, D. (2017). Big Data and the Future of R&D Management. *Research Technology Management*, *60*(5), 43–50. DOI:10.1080/08956308.2017.1348135
- Boaz, A., Ashby, D., Denyer, D., Egan, M., Harden, A., Jones, D. R., Pawson, R., & Tranfield, D. (2006). A multitude of syntheses: A comparison of five approaches from diverse policy fields. *Evidence & Policy*, *2*(4), 479–502. DOI:10.1332/174426406778881755
- Bresciani, S., Ciampi, F., Meli, F., & Ferraris, A. (2021). Using Big Data for Co-Innovation Processes: Mapping the Field of Data-driven Innovation, Proposing theoretical Developments and Providing a Research Agenda. *International Journal of Information Management*, *60*. DOI:10.1016/j.ijinfomgt.2021.102347
- Buganza, T., Trabucchi, D., & Pellizzoni, E. (2020). Limitless personalisation: The role of Big Data in unveiling service opportunities. *Technology Analysis & Strategic Management*, *32*(1), 58–70. DOI:10.1080/09537325.2019.1634252
- Campbell-Hunt, C. (2000). What Have We Learned About Generic Competitive Strategy? A Meta-Analysis. *Strategic Management Journal*, *21*(2), 127–154. DOI: 10.1002/(SICI)1097-0266(200002)21:2<127::AID-SMJ75>3.0.CO;2-1
- Cather, D. A. (2018). Cream Skimming: Innovations in Insurance Risk Classification and Adverse Selection. *Risk Management and Insurance Review*, *21*(2), 335–366. DOI:10.1111/rmir.12102
- Carr, N.G. (2003). IT Doesn't Matter. *Harvard Business Review*, *81*(5), 41–49.
- Chae, B. (Kevin). (2019). A General framework for Studying the Evolution of the Digital Innovation Ecosystem: The case of big data. *International Journal of Information Management*, *45*, 83–94. DOI:10.1016/j.ijinfomgt.2018.10.023
- Chai, S., & Shih, W. (2017). Why Big Data Isn't Enough. *MIT Sloan Management Review*, *58*(2), 57–61.
- Chandy, R., Hassan, M., & Mukherji, P. (2017). Big Data for Good: Insights from Emerging Markets. *Journal of Product Innovation Management*, *34*(5), 703–713. DOI:10.1111/jpim.12406

- Chen, C. L. P., & Zhang, C.-Y. (2014). Data-intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data. *Information Sciences*, 275, 314–347. DOI:10.1016/j.ins.2014.01.015
- Chen, Jie. (2019). The Influence of Big Data on Transformation of Traditional Manufacturing Enterprises from the Perspective of Business Model. *Management & Engineering*, 34, 3–8. DOI:10.5503/J.ME.2019.34.001
- Chen, Jinchuan, Chen, Y., Du, X., Li, C., Lu, J., Zhao, S., & Zhou, X. (2013). Big data Challenge: A Data Management Perspective. *Frontiers of Computer Science*, 7(2), 157–164. DOI:10.1007/s11704-013-3903-7
- Chen, P. T. (2018). Medical Big Data Applications: Intertwined Effects and Effective Resource Allocation Strategies identified through IRA-NRM Analysis. *Technological Forecasting and Social Change*, 130, 150–164. DOI:10.1016/j.techfore.2018.01.033
- Chesbrough, H. W. (2003). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Cambridge: Harvard Business School Publishing.
- 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*, 24(5), 1157–1176. DOI:10.1108/JKM-02-2020-0156
- Cohen, W.M. & Levinthal, D.A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), 128–152. DOI: 10.2307/2393553
- Coussement, K., Debaere, S., & De Ruyck, T. (2017). Inferior Member Participation Identification in Innovation Communities: The Signaling Role of Linguistic Style Use. *Journal of Product Innovation Management*, 34(5), 565–579. DOI:10.1111/jpim.12401
- Crossan, M. M., & Apaydin, M. (2010). A Multi-Dimensional Framework of Organizational Innovation: A Systematic Review of the Literature. *Journal of Management Studies*, 47(6), 1154–1191. DOI:10.1111/j.1467-6486.2009.00880.x
- Dougherty, D., Bertels, H., Chung, K., Dunne, D. D., & Kraemer, J. (2013). Whose Time Is It? Understanding Clock-time Pacing and Event-time Pacing in Complex Innovations. *Management and Organization Review*, 9(2), 233–263. DOI:10.1111/more.12017
- De Luca, L. M., Herhausen, D., Troilo, G., & Rossi, A. (2021). How and when do big data investments pay off? The role of marketing affordances and service innovation. *Journal of the Academy of Marketing Science*, 49(4), 790–810. DOI:10.1007/s11747-020-00739-x
- De Smet, A., & Gagnon, C. (2018). Organizing for the Age of Urgency. *McKinsey Quarterly*, 1, 1–11.
- Del Vecchio, P., Di Minin, A., Petruzzelli, A. M., Panniello, U., & Pirri, S. (2018). Big Data for Open Innovation in SMEs and Large Corporations: Trends, Opportunities, and Challenges. *Creativity and Innovation Management*, 27(1), 6–22. DOI:10.1111/caim.12224
- Demchenko, Y., Grosso, P., De Laat, C., & Membrey, P. (2013). Addressing Big Data Issues in Scientific Data Infrastructure. *Proceedings of the 2013 International Conference on Collaboration Technologies and Systems*, 48–55. DOI:10.1109/CTS.2013.6567203
- Dicuonzo, G., Galeone, G., Zappimulso, E., & Dell'Atti, V. (2019). Risk Management 4.0: The Role of Big Data Analytics in the Bank Sector. *International Journal of Economics and Financial Issues*, 9(6), 40–47.

- Dubé, L., Du, P., McRae, C., Sharma, N., Jayaraman, S., & Nie, J.-Y. (2018). Convergent Innovation in Food through Big Data and Artificial Intelligence for Societal-Scale Inclusive Growth. *Technology Innovation Management Review*, 8(2), 49–65.
- Duncan, E., Alesandros, G., Ross, D. Z., & Eric, N. (2021). New but for whom? Discourses of Innovation in Precision Agriculture. *Agriculture and Human Values*, 38(4), 1181–1199. DOI:10.1007/s10460-021-10244-8
- El-Kassar, A.-N., & Singh, S. K. (2019). Green Innovation and Organizational Performance: The influence of Big Data and the Moderating Role of Management Commitment and HR Practices. *Technological Forecasting and Social Change*, 144, 483–499. DOI:10.1016/j.techfore.2017.12.016
- Etzkowitz, H. (2008). *The Triple Helix. University-Industry-Government Innovation in Action*. New York: Routledge.
- Fink, A. (2013). *Conducting Research Literature Reviews: From the Internet to Paper* (4th ed.). Los Angeles: SAGE Publications.
- Fleckenstein, M., & Fellows, L. (2018). Data Analytics. In M. Fleckenstein & L. Fellows (Eds.), *Modern Data Strategy*, (pp. 133-142). Springer International Publishing. DOI:10.1007/978-3-319-68993-7_13
- Fortunato, A., Gorgoglione, M., Messeni Petruzzelli, A., & Panniello, U. (2017). Leveraging Big Data for Sustaining Open Innovation: The Case of Social TV. *Information Systems Management*, 34(3), 238–249. DOI:10.1080/10580530.2017.1330000
- Freeman, L.C. (1978). Centrality in Social Networks Conceptual Clarification. *Social Networks*, 1(3), 215–239. DOI: 10.1016/0378-8733(78)90021-7
- Franke, N., von Hippel, E. & Schreier, M. (2006). Finding Commercially Attractive User Innovations: A Test of Lead-User Theory. *Journal of Product Innovation Management*, 23(4), 301–315. DOI: 10.1111/j.1540-5885.2006.00203.x
- Frizzo-Barker, J., & Chow-White, P. A. (2014). From Patients to Petabytes: Genomic Big Data, Privacy, and Informational Risk. *Canadian Journal of Communication*, 39(4), 615–625.
- Gao, J., Li, W., Zhao, Z., & Han, Y. (2020). Provisioning Big Data Applications as Services on Containerised Cloud: A Microservices-Based Approach. *International Journal of Services Technology and Management*, 26(2–3), 167–181. DOI:10.1504/IJSTM.2020.106744
- Gopalakrishnan, S. & Damanpour, F. (1997). A Review of Innovation Research in Economics, Sociology and Technology Management. *Omega*, 25(1), 15–28. DOI: 10.1016/S0305-0483(96)00043-6
- Gradeck, J., Neguriță, O., Grecu, I., & Grecu, G. (2019). Big Data Analytics in Industry 4.0: Sustainable Industrial Value Creation, Manufacturing Process Innovation, and Networked Production Structures. *Journal of Self-Governance and Management Economics*, 7(3), 34–40. DOI:10.22381/JSME7320195
- Graves, A. (1987). Comparative Trends in Automotive Research and Development. DRC Discussion Paper No. 54. Brighton: Sussex University
- Grimpe, C., Sofka, W., Bhargava, M. & Chatterjee, R. (2017). R&D, Marketing Innovation, and New Product Performance: A Mixed Methods Study. *Journal of Product Innovation Management*, 34(3), 360–383. DOI: 10.1111/jpim.12366

- Gusenbauer, M. (2022). Search where you will find most: Comparing the disciplinary coverage of 56 bibliographic databases. *Scientometrics*, *127*(5), 2683–2745. DOI:10.1007/s11192-022-04289-7
- Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, *11*(2), 181–217. DOI:10.1002/jrsm.1378
- Hambrick, D.C. & Mason, P.A. (1984). Upper Echelons: The Organization as a Reflection of Its Top Managers. *The Academy of Management Review*, *9*(2), 193–206. DOI: 10.2307/258434
- Hobday, M. (2005). Firm-level Innovation Models: Perspectives on Research in Developed and Developing Countries. *Technology Analysis and Strategic Management*, *17*(2), 121–146. DOI:10.1080/09537320500088666
- Hodgkinson, G.P. & Ford, J.K. (2014). Narrative, Meta-Analytic, and Systematic Reviews: What Are the Differences and Why Do They Matter? *Journal of Organizational Behavior*, *35*(S1), S1–S5. DOI: 10.1002/job.1918
- Huang, Y., Porter, A. L., Cunningham, S. W., Robinson, D. K. R., Liu, J., & Zhu, D. (2018). A technology delivery system for characterizing the supply side of technology emergence: Illustrated for Big Data & Analytics. *Technological Forecasting and Social Change*, *130*, 165–176. DOI:10.1016/j.techfore.2017.09.012
- Iorio, C., Pandolfo, G., D'Ambrosio, A., & Siciliano, R. (2020). Mining Big Data in Tourism. *Quality and Quantity*, *54*(5–6), 1655–1669. DOI:10.1007/s11135-019-00927-0
- Johnson, J. S., Friend, S. B., & Lee, H. S. (2017). Big Data Facilitation, Utilization, and Monetization: Exploring the 3Vs in a New Product Development Process. *Journal of Product Innovation Management*, *34*(5), 640–658. DOI:10.1111/jpim.12397
- Kayser, V., Nehrke, B., & Zubovic, D. (2018). Data Science as an Innovation Challenge: From Big Data to Value Proposition. *Technology Innovation Management Review*, *8*(3), 16–25.
- Kim, J. Y., & Lee, M. J. (2016). Living with casinos: The Triple-Helix Approach, Innovative Solutions, and Big Data. *Technological Forecasting and Social Change*, *110*, 33–41. DOI:10.1016/j.techfore.2016.04.024
- Kline, S. J., & Rosenberg, N. (1986). An Overview of Innovation. In R. Landau & N. Rosenberg (Eds.), *The Positive Sum Strategy. Harnessing Technology for Economic Growth*, (pp. 275–305). Washington: National Academy Press. DOI:10.1142/9789814273596_0009
- Kozak, J., Kania, K., Juszczuk, P., & Mitreęga, M. (2021). Swarm intelligence goal-oriented approach to data-driven innovation in customer churn management. *International Journal of Information Management*, *60*, 1–16. DOI:10.1016/j.ijinfomgt.2021.102357
- Lambrou, M. (2016). Innovation Capability, Knowledge Management and Big Data Technology: A Maritime Business Case. *International Journal of Advanced Corporate Learning*, *9*(2), 40. DOI:10.3991/ijac.v9i2.6010
- Langevin, M. (2019). Big Data for (not so) Small Loans: Technological Infrastructures and the Massification of Fringe Finance. *Review of International Political Economy*, *26*(5), 790–814. DOI:10.1080/09692290.2019.1616597
- Lee, H. L. (2018). Big Data and the Innovation Cycle. *Production and Operations Management*, *27*(9), 1642–1646. DOI:10.1111/poms.12845

- Lieberman, M.B. & Montgomery, D.B. (1988). First-Mover Advantages. *Strategic Management Journal*, 9(S1), 41–58. DOI: 10.1002/smj.4250090706
- Lim, K. Y. H., Zheng, P., & Chen, C. H. (2020). A state-of-the-art survey of Digital Twin: Techniques, Engineering Product Lifecycle Management and Business Innovation Perspectives. *Journal of Intelligent Manufacturing*, 31(6), 1313–1337. DOI:10.1007/s10845-019-01512-w
- Lin, R., Xie, Z., Hao, Y., & Wang, J. (2020). Improving High-Tech Enterprise Innovation in Big Data Environment: A Combinative View of Internal and External Governance. *International Journal of Information Management*, 50, 575–585. DOI:10.1016/j.ijinfomgt.2018.11.009
- Liu, H., & Liu, X. (2019). A Novel Research on the Influence of Enterprise Culture on Internal Control in Big Data and Internet of Things. *Mobile Networks and Applications*, 24(2), 365–374. DOI:10.1007/s11036-017-0910-1
- Liu, Y., Soroka, A., Han, L., Jin, J., & Tang, M. (2020). Cloud-based Big Data Analytics for Customer Insight-Driven Design Innovation in SMEs. *International Journal of Information Management*, 51, 1–12. DOI:10.1016/j.ijinfomgt.2019.11.002
- Liu, Z., Ma, L., Zhu, Y., & Ji, W. (2019). An Investigation on Responsible Innovation in the Emerging Shared Bicycle Industry: Case Study of a Chinese Firm. *Journal of Open Innovation*, 5(3), 42. DOI:10.3390/joitmc5030042
- Ma, X., & Zhang, Y. (2021). Digital Innovation Risk Management Model of Discrete Manufacturing Enterprise Based on Big Data Analysis. *Journal of Global Information Management*, 30(7), 1–14. DOI:10.4018/JGIM.286761
- Mann, L. (2018). Left to Other Peoples' Devices? A Political Economy Perspective on the Big Data Revolution in Development. *Development and Change*, 49(1), 3–36. DOI:10.1111/dech.12347
- Markides, C. & Charitou, C.D. (2004). Competing with Dual Business Models: A Contingency Approach. *Academy of Management Perspectives*, 18(3), 22–36. DOI: 10.5465/ame.2004.14776164
- Masih, J., & Joshi, A. (2021). Understanding Health-Foods Consumer Perception Using Big Data Analytics. *Journal of Management Information and Decision Sciences*, 24(S3), 1–15.
- Mastrangelo, D. G. (2018). Connectedness: Enhancing the Innovative Capacity of an Organization in an Era of Big Data. *Prestige International Journal of Management and Research*, 11R(1–2), 50–57.
- McCain, R.A. (2010). *Game Theory: A Nontechnical Introduction to the Analysis of Strategy (3rd edition)*. Hackensack, NJ: World Scientific Publishing.
- McGregor, C. A. M. (2018). Using Constructive Research to Structure the Path to Transdisciplinary Innovation and Its Application for Precision Public Health with Big Data Analytics. *Technology Innovation Management Review*, 8(8), 7–15.
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment. *British Journal of Management*, 30(2), 272–298. DOI:10.1111/1467-8551.12343
- Moore, G.A. (1999). *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*. New York: Harper Collins Publishers.
- Nayar, R., & Nayar, S. (2014). Big Data and Leadership for Innovation - An Aggregator Model. *Journal of Applied Management*, 6(1), 109–117.

- Nuccio, M., & Guerzoni, M. (2019). Big data: Hell or heaven? Digital platforms and market power in the data-driven economy. *Competition & Change*, 23(3), 312–328. DOI:10.1177/1024529418816525
- Nunan, D., & Di Domenico, M. (2017). Big Data: A Normal Accident Waiting to Happen? *Journal of Business Ethics*, 145(3), 481–491. DOI:10.1007/s10551-015-2904-x
- OECD (2015). *Data-Driven Innovation: Big Data for Growth and Well-Being*. Paris: OECD Publishing. DOI:10.1787/9789264229358-en
- OECD & Eurostat (2018). *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation (4th ed.)*. Paris: OECD Publishing/Luxembourg: Eurostat. DOI:10.1787/24132764
- Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: A handbook for visionaries, game changers, and challengers*. Hoboken, NJ: Wiley.
- Prescott, M. E. (2016). Big Data: Innovation and Competitive Advantage in an Information Media Analytics Company. *Journal of Innovation Management*, 4(1), 92–113.
- Press, G. (2018). *Why Facebook Triumphed Over All Other Social Networks*. Forbes, April 8, 2018. www.forbes.com/sites/gilpress/2018/04/08/why-facebook-triumphed-over-all-other-social-networks. Last accessed December 18, 2022.
- Pricop, A. G., & Orzan, G. (2018). Big Data: The Ubiquitous Data That Organizations Need to Leverage Nowadays. *Annales Universitatis Apulensis Series Oeconomica*, 20(1), 15–20.
- Rejeb, A., Rejeb, K., & Keogh, J. G. (2020). Potential of Big Data for Marketing: A Literature Review. *Management Research and Practice*, 12(3), 60–73.
- Rodriguez, L., & Da Cunha, C. (2018). Impacts of Big data Analytics and Absorptive Capacity on Sustainable Supply Chain Innovation: A Conceptual Framework. *LogForum*, 14(2). DOI:10.17270/J.LOG.267
- Rogers, E.M. (2003). *Diffusion of Innovations (5th ed.)*. New York, London, Toronto, Sydney: Free Press.
- Rothwell, R. (1994). Towards the Fifth-Generation Innovation Process. *International Marketing Review*, 11(1), 7–31. DOI:10.1108/02651339410057491
- Rousseau, D. M., Manning, J., & Denyer, D. (2008). Evidence in Management and Organizational Science: Assembling the Field's Full Weight of Scientific Knowledge Through Syntheses. *Academy of Management Annals*, 2(1), 475–515. DOI:10.5465/19416520802211651
- Saggi, M. K., & Jain, S. (2018). A Survey Towards an Integration of Big Data Analytics to Big Insights for Value-Creation. *Information Processing & Management*, 54(5), 758–790. DOI:10.1016/j.ipm.2018.01.010
- Saheb, T., & Saheb, T. (2020). Understanding the Development Trends of Big Data Technologies: An Analysis of Patents and the Cited Scholarly Works. *Journal of Big Data*, 7(12). DOI:10.1186/s40537-020-00287-9
- Saldžiūnas, K., & Skyrius, R. (2017). The Challenges of Big Data Analytics in the Mobile Communications Sector. *Ekonomika*, 96(2), 110–121. DOI:10.15388/Ekon.2017.2.11004
- Sati, Z. E. (2017). Evaluation of Big Data and Innovation Interaction in Increase Supply Chain Competencies. *Universitpark Bülten*, 6(2), 88–102. DOI:10.22521/unibulletin.2017.62.7

- Savino, T., Messeni Petruzzelli, A. & Albino, V. (2017). Search and Recombination Process to Innovate: A Review of the Empirical Evidence and a Research Agenda. *International Journal of Management Reviews*, 19(1), 54–75. DOI: 10.1111/ijmr.12081
- Schmookler, J. (1966). *Invention and Economic Growth*. University of Michigan: Harvard University Press.
- Scott, W.R. (2014). *Institutions and Organizations: Ideas, Interests, and Identities*. Los Angeles, CA: SAGE.
- Scuotto, V., Santoro, G., Bresciani, S., & Del Giudice, M. (2017). Shifting Intra- and Inter-Organizational Innovation Processes Towards Digital Business: An Empirical Analysis of SMEs. *Creativity and Innovation Management*, 26(3), 247–255. DOI:10.1111/caim.12221
- Seggie, S. H., Soyer, E., & Pauwels, K. H. (2017). Combining Big Data and Lean Startup Methods for Business Model Evolution. *AMS Review*, 7(3), 154–169. DOI:10.1007/s13162-017-0104-9
- Shankar, V., Carpenter, G.S. & Krishnamurthi, L. (1998). Late Mover Advantage: How Innovative Late Entrants Outsell Pioneers. *Journal of Marketing Research*, 35(1), 54–70. DOI: 10.1177/002224379803500107
- Sollins, K. R. (2019). IoT Big Data Security and Privacy Versus Innovation. *IEEE Internet of Things Journal*, 6(2), 1628–1635. DOI:10.1109/JIOT.2019.2898113
- Somoza Sanchez, V. V., Giacalone, D., & Goduscheit, R. C. (2018). Digital Anthropology as Method for Lead User Identification from Unstructured Big Data. *Creativity and Innovation Management*, 27(1), 32–41. DOI:10.1111/caim.12228
- Sonka, S. (2016). Big Data: Fueling the Next Evolution of Agricultural Innovation. *Journal of Innovation Management*, 4(1), 114–136.
- Sorescu, A. (2017). Data-Driven Business Model Innovation. *Journal of Product Innovation Management*, 34(5), 691–696. DOI:10.1111/jpim.12398
- Steinberg, P. J., Asad, S., & Lijzenga, G. (2022). Narcissistic CEOs' dilemma: The trade-off between exploration and exploitation and the moderating role of performance feedback. *Journal of Product Innovation Management*, 39(6), 773–796. DOI: 10.1111/jpim.12644
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18(7), 509–533. DOI: 10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z
- Tidd, J., Bessant, J. R., & Pavitt, K. (2005). *Managing Innovation: Integrating Technological, Market and Organization Change (3rd ed.)*. Chichester: John Wiley & Sons.
- Tornatzky, L.G. & Fleischer, M. (1990). *The Processes of Technological Innovation (3rd ed.)*. Lexington, MA: Lexington Books.
- Trabucchi, D., Buganza, T., Dell'Era, C., & Pellizzoni, E. (2018). Exploring the inbound and outbound strategies enabled by user generated big data: Evidence from leading smartphone applications. *Creativity and Innovation Management*, 27(1), 42–55. DOI:10.1111/caim.12241
- Trabucchi, D., & Buganza, T. (2019). Data-Driven Innovation: Switching the Perspective on Big Data. *European Journal of Innovation Management*, 22(1), 23–40. DOI:10.1108/EJIM-01-2018-0017

- Trabucchi, D., & Buganza, T. (2020). Fostering digital platform innovation: From two to multi-sided platforms. *Creativity and Innovation Management*, 29(2), 345–358. DOI:10.1111/caim.12320
- Trabucchi, D., & Buganza, T. (2022). Landlords with No Lands: A Systematic Literature Review on Hybrid Multi-Sided Platforms and Platform Thinking. *European Journal of Innovation Management*, 25(6), 64–96. DOI:10.1108/EJIM-11-2020-0467
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14, 207–222. DOI: 10.1111/1467-8551.00375
- Troilo, G., De Luca, L. M., & Guenzi, P. (2017). Linking Data-Rich Environments with Service Innovation in Incumbent Firms: A Conceptual Framework and Research Propositions. *Journal of Product Innovation Management*, 34(5), 617–639. DOI:10.1111/jpim.12395
- Tushman, M.L. & O'Reilly, C.A. (1996). Ambidextrous Organizations: Managing Evolutionary and Revolutionary Change. *California Management Review*, 38 (4), 8–29. DOI: 10.2307/41165852
- van den Broek, T., & van Veenstra, A. F. (2018). Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. *Technological Forecasting and Social Change*, 129, 330–338.
- van de Ven, A.H. & Poole, M.S. (1995). Explaining Development and Change in Organizations. *Academy of Management Review*, 20(3), 510–540. DOI: 10.5465/amr.1995.9508080329
- von Grafenstein, M., Wernick, A., & Olk, C. (2019). Data Governance: Enhancing Innovation and Protecting Against Its Risks. *Intereconomics*, 54(4), 228–232. DOI:10.1007/s10272-019-0829-9
- von Hippel, E. (1986). Lead Users: A Source of Novel Product Concepts. *Management Science*, 32(7), 791–805. DOI:10.1287/mnsc.32.7.791
- Ward, J. S., & Barker, A. (2013). Undefined By Data: A Survey of Big Data Definitions. *arXiv* 1309.5821.
- West, J., & Bogers, M. (2014). Leveraging External Sources of Innovation: A Review of Research on Open Innovation. *Journal of Product Innovation Management*, 31(4), 814–831. DOI:10.1111/jpim.12125
- Wright, L. T., Robin, R., Stone, M., & Aravopoulou, E. (2019). Adoption of Big Data Technology for Innovation in B2B Marketing. *Journal of Business to Business Marketing*, 26(3–4), 281–293. DOI:10.1080/1051712X.2019.1611082
- Xu, F., Nash, N., & Whitmarsh, L. (2020). Big Data or Small Data? A Methodological Review of Sustainable Tourism. *Journal of Sustainable Tourism*, 28(2), 144–163. DOI:10.1080/09669582.2019.1631318
- Yablonsky, S. A. (2019). Multidimensional Data-Driven Artificial Intelligence Innovation. *Technology Innovation Management Review*, 9(12), 16–28.
- Yamamoto, K., & Lloyd, R. A. (2019). The Role of Big Data and Digitization in Just-In-Time (JIT) Information Feeding and Marketing. *American Journal of Management*, 19(2), 126–133. DOI:10.33423/ajm.v19i2.2074
- Yin, R. K. (2014). *Case Study Research: Design and Methods* (5th edition). Los Angeles: SAGE.

Zhang, Y., Huang, Y., Porter, A. L., Zhang, G., & Lu, J. (2019). Discovering and forecasting interactions in big data research: A learning-enhanced bibliometric study. *Technological Forecasting and Social Change*, 146, 795–807. DOI:10.1016/j.techfore.2018.06.007

Biographies



Nebojša Radojević. Nebojša Radojević is a Professor of Project Management at IU International University of Applied Sciences in Dresden, Germany. Previously, he was a Professor of Strategic Management and Project Management at Shenzhen Technology University in Shenzhen, China. His research primarily elaborates on the innovation process at the firm level in a broader technological, societal and managerial context. Before his career in academia, he accumulated more than 20 years of professional experience as manager, project manager and management consultant for multinational corporations and higher levels of government in Germany, Canada and Cyprus. He holds a PhD in International Business and an MBA from University of Montreal, a master's degree in Computer Science from Dresden University of Technology, and professional credentials in project and program management.

ORCID: <https://orcid.org/0009-0008-8160-870X>

CRedit Statement: *Conceptualization, Data Curation, Formal Analysis, Investigation, Visualization, Methodology, Writing – Original Draft, Writing – Review & Editing, Supervision.*



Erich Heumüller. Erich Heumüller is a Professor for Business Information Systems and the director of the Center for Digital Transformation at Baden-Wuerttemberg Cooperative State University in Stuttgart, Germany. In his research, he focusses on digital transformation, digital ecosystems, and digital technologies. Before his career in academia, he accumulated approx. 10 years of professional experience as army officer and management consultant for multinational corporations and higher levels of government in Germany. He received a PhD and a Diploma degree in Information Systems from the University of the Armed Forces Munich, and holds professional credentials in IT governance, IT service management, and project management.

ORCID: <https://orcid.org/0000-0003-3479-4067>

CRedit Statement: *Conceptualization, Data Curation, Formal analysis, Investigation, Visualization, Methodology, Writing – Original Draft, Project administration.*