Journal of Scientometric Research, 2024; 13(3):791-805.
https://www.jscires.org

Research Article

Unveiling the Evolution of Big Data Analytics Capability: A Comprehensive Analysis

Sunil Pathak[1,*], Venkataraghavan Krishnaswamy[2], Mayank Sharma[2]

[1]NMIMS Centre for Distance and Online Education, SVKM NMIMS, Mumbai, Maharashtra, INDIA.
[2]Information Technology and Systems Area, Indian Institute of Management Kashipur, Kashipur, Uttarakhand, INDIA.

ABSTRACT

This paper performs a state-of-the-art literature review of Big Data Analytics Capability (BDAC) to analyze its domain, knowledge structures, theoretical roots, and key research trends. We analyze the domain, knowledge structures, and evolution of the field by applying Lotka's law, Bradford law, MCP ratio, keyword, citation, and co-citation analysis. Our analysis reveals predominant use of resource-based view and dynamic capability theories among other theories used in BDAC research. The key research themes identified relate to BDAC antecedents, consequences, process/industry contexts, business intelligence, advanced analytics, and environment dynamics. The study's implications pertain to the identification of BDAC development, its theoretical roots, and emerging research themes. We also identify interesting research opportunities relating to BDAC as a dynamic capability, BDAC challenges such as failures, maturity, response to market dynamics, and BDAC value at process, firm, and industry levels.

Keywords: Big Data, Big Data Analytics, Big Data Analytics Capability, Evolution, Literature Review, Dynamic Capability.

Correspondence:
Sunil Pathak
NMIMS Centre for Distance and Online Education, SVKM NMIMS, Mumbai, Maharashtra, INDIA.
Email: sunil.pathak.web@gmail.com
ORCID: 0000-0003-4872-4596

Received: 06-01-2024;
Revised: 11-05-2024;
Accepted: 02-08-2024.

INTRODUCTION

The emergence of Big Data as a disruptive paradigm in recent years can be attributed to the substantial growth of data stemming from advancements in technology, including social media, mobile devices, cloud computing, and the IoT. Recognizing its strategic significance, organizations have made significant investments in the development of infrastructure and capabilities tailored for Big Data over the past few years.[1-4] However, despite its potential, the distinct challenges posed by Big Data-namely, its volume, velocity, variety, and veracity-coupled with intricate information value chains and digitized business environments, have challenged organizations from fully capitalizing on Big data capabilities due to deficiencies in requisite skill sets, resources, and capabilities.[5,6] Within this context, the emerging multidimensional concept of Big Data Analytics Capability (BDAC) has garnered significant attention. BDAC encompasses an organization's capacity to effectively harness Big Data, facilitated by the alignment of people, processes, and technologies, for the purpose of creating strategic and operational value.[7,8] Early scholarly endeavors focused on elucidating the ontology of BDAC,[9] while recent studies have focused on the mechanisms underlying business value creation within the context of BDAC.[10-12]

Despite the considerable business potential of big data, it is anticipated that only a modest 20% of endeavors will yield discernible business outcomes.[6] Consequently, there exists a pressing need for regular scholarly inquiry for a comprehensive view of the historical, contemporary, and prospective landscape of BDAC research. Despite the acknowledged significance of Big Data Analytics (BDA), BDAC research remains nascent, with a focus on the technical dimensions of BDA, big data technologies and their applications. There is a noticeable paucity of inquiries addressing the broader capabilities of BDAC and a dearth of research around managerial issues.[8] For instance, Phillips-Wren *et al.*[13] provided early direction for BDA from a data collection, storage, and processing perspective. Singh,[14] on the other hand, reviewed a narrow span of literature between 2015-18 and identified engineering and deployment issues concerning big data technologies such as visualization, data mining, analytics, and data preparation tools. Qazi and Sher[15] and Chong and Shi[16] discussed the scope and applications of BDA. Similarly, Khanra *et al.*,[8] in their bibliometric analysis, identified key research themes for big data analytics, analytical tools, management domains, and ethical issues. While these studies suggest the growing and continued research interest in big data, the knowledge about BDAC construct, underlying theoretical framework,

OPEN ACCESS

Manuscript TECHNOMEDIA

DOI: 10.5530/jscires.20041105

Publishing Partner : Manuscript Technomedia. [www.mstechnomedia.com]

research progression, and management challenges needs further investigation.[17]

Within the ambit of BDAC, Mikalef *et al.*,[10] through a systematic literature review, proposed a conceptual framework delineating the business value of BDAC, whereas Pathak *et al.*[18] introduced an integrated framework that amalgamates BDAC, Dynamic Capability Framework (DCF) with enterprise architecture. These contributions underscore the imperative for further scholarship aimed at elucidating resource orchestration, trust dynamics, business value measurement, and the decision-making processes facilitated by BDAC.

In this paper, unlike other studies, we focus only on BDAC as the subject of inquiry, which is hitherto discussed sparsely covering a broad horizon of sixteen years from 2007. We follow a bibliometric centric approach to uncover the domain, knowledge structures, theoretical roots, and key research trends. We have chosen bibliometrics analysis as it provides a means to analyze the literature by focusing on the domain and knowledge structures of a specific field. The domain analysis focuses on different units of analysis such as sources, authors and documents, and associated metrics. A knowledge structure, on the other hand, is an interrelated collection of knowledge about a particular topic related to the conceptual (using co-word analysis), intellectual (co-citation), or social (collaboration) roots of a field. Such analysis helps to uncover theories, objectives, and emerging research themes in a topical field such as BDAC and is preferred over pure bibliometric reviews and is much useful to the research community.[19] Therefore, our work differs from the existing reviews in both the objectives, currency, and the methodology to address the following questions:

- What is the maturity of BDAC research?
- How has the BDAC research evolved?
- What are the key research trends and opportunities?

The study makes two primary contributions: firstly, it identifies and discusses the developmental trajectory and theoretical underpinnings of Big Data Analytics Capability (BDAC) phenomena, alongside emerging research themes; secondly, it delineates research opportunities pertaining to BDAC as a Dynamic Capability (DC), encompassing challenges such as failures, maturity, responsiveness to market dynamics, and value assessment at process, firm, and industry levels. The paper is structured as follows: the subsequent section furnishes a background on BDAC, succeeded by an exposition of the methodologies employed for data collection. Subsequently, an analysis of domain and knowledge structures is presented, followed by a discussion of key themes. The concluding section elucidates the implications, research opportunities, limitations, and overall contribution of the study.

BDAC- A Background

The emerging digital and technological innovations are a source of massive disruption and opportunities to any modern enterprise.[20] As competition moves from a physical plane to a virtual plane, these disruptions are evident in rising consumer expectations, the fast-changing competitive landscape, and the massive amounts of data generated from digital technologies.[21] Consequently, the traditional IS strategy's functional role needs to be transformed to an overarching role suggesting a fusion between IS and business strategies.[1] Authors defined this integrated role as the "strategy formulated and executed by leveraging digital resources to create differential value". Big data is at the center of such phenomena and a part of a firm's digital and strategic toolkit. Firms view big data as a business asset and an enabler.[22,23] which has the potential to change the market dynamics by creating new data-centric products/services (such as Spotify or LinkedIn APIs), creating new industries around data tools (such as snowflake), or monetizing the data (such as APIs published by Meta or LinkedIn), thus impacting the entire business value chain. These examples highlight the need for specialized capabilities such as BDAC to derive business value from big data.

Literature on Big Data reveals a focus on concepts such as BDA and BDAC. BDA has been commonly referred to as the tools, technologies, methods, and processes to manage and exploit Big Data.[24,25] On the other hand, BDAC is described as the capability to exploit Big Data to generate strategic and operational value.[7,26] The proliferation of Big Data has raised management concerns regarding its development, adoption, use, regulatory challenges, liability, privacy, ethics, and security. While BDAC research is still nascent, significant advances were made within academia in the last 10 years. However, the transition from an information system-centric capability to a digital/strategic capability is still surrounded by uncertainties concerning the value creation mechanisms and the complex interplay of technology and people, organizational routines, business processes, and business models.[1,10-12] Pathak *et al.*[18] established the inconsistencies in BDAC definitions and called for integrating existing views. These studies highlight that a thorough investigation of the existing research is required to synthesize the knowledge on BDAC.

METHODOLOGY

Database and Search Protocol

This study followed a multi-step protocol comprising of database search, analysis, and findings. We used several tools and techniques to analyze domain and knowledge structures using descriptive analysis, citation, and co-citation analysis. We used the Scopus database for document search,[27] VOSviewer[28] and Bibliometrix R package[29] for analysis and visualization.

We extracted the peer-reviewed articles from SCOPUS using the following search string-'TITLE-ABS-KEY ("BDAC" OR "Big Data Analytics Capability" OR "Analytics Capability" OR "BI. Capability "OR "Big Data Capability" OR "Business Intelligence Capability" OR "Big Data Organization Capability " OR "Big Data Analytics Factors" OR "Big Data Dynamic Capability" OR "Big Data Management Capability".

The initial search resulted in 753 documents and was manually verified for relevance by examining their titles. Several documents related to the computer science and engineering aspects of BDAC were dropped in this step, reducing the corpus to 352 documents. In the next step, we reviewed the 352 abstracts and keywords for relevance and pruned the collection to 162 documents. For bibliometric reviews, Rogers *et al.*[30] recommend an analytical minimum of fewer than 200 papers and justify even smaller samples for qualitative and concept-based reviews for emerging fields such as BDAC. Such concept-based bibliometric research of an emerging field such as BDAC, offers greater focus, in-depth exploration, precision, reduces noise and complexity (typically for large data sets), and interpretability leading to more meaningful and interpretable findings.

Table 1 summarizes the dataset statistics. The dataset consisted of 162 articles between year 2007-2013 which had BDAC as the focus concept from 104 publications (or sources). In this work, a citing document refers to a research article (such as a journal article or conference proceeding) included in a bibliographic collection. A cited reference refers to a scientific document included in at least one of the reference lists of the bibliographic collection. We found 11094 cited references in our bibliographic collection of 162 documents.

Analysis

A detailed bibliometric analysis of a specific subject requires analysis of document information from its metadata such as keywords, authors, references, sources, countries, and future research directions.[31] It is often defined as "the quantitative analysis of publications in a given field" aimed at identifying its structure and development.[32] This paper analyses the dataset in two parts. The first part, the descriptive analysis, focuses on the domain aspects and delineates the field of the study. The second part performs an in-depth analysis of the knowledge structures to uncover the field's evolution and derive the main research trends.

Descriptive Analysis

The descriptive analysis aims to uncover the domain aspects of the topic, such as the number of documents published per year, documents by subject area, distribution of documents in sources and publications, key authors, and scientific productivity. We used the Bibliometrix R package and biblioshiny application[29] for descriptive analysis tasks. The identified documents were extracted from Scopus in BibTeX format and were then imported into biblioshiny. Finally, the dataset was analyzed for specific metrics such as dispersion,[33] scientific productivity by Lotka, 1926[34] and MCP ratio.

Knowledge Structures

This step covered the knowledge aspects to generate the conceptual and intellectual structure of the field. The specific bibliographic techniques used in this step involved keyword, citation, and co-citation analysis through VOSviewer.[28] First, the keyword analysis helps in uncovering the prominent themes and research clusters. Next, the intellectual structure is identified by the citation and co-citation analysis. While citation analysis focuses on the cited documents included in the research dataset, co-citation analysis relies upon the references cited by the documents. We discuss the analysis and findings in forthcoming sections.

Descriptive Analysis

Documents

The publication growth trends during this period are shown in Figure 1. The data (total of 162 documents) indicates that while the growth was slow between 2007 and 2015, it grew rapidly in the last five years (2016-23). This surge in growth serves as a testament to the burgeoning status of BDAC as an emergent phenomenon that has captured substantial attention within academic research circles. Table 2 provides an overview of publication distribution across various countries, with the United States emerging as the foremost contributor, followed by China, Australia, the United Kingdom, and India. Notably, a considerable proportion of these publications are situated within the domain of business and management research, reflecting a concerted effort to contextualize BDAC within socio-technical frameworks. This underscored emphasis on interdisciplinary engagement highlights the pervasive influence of BDA, transcending

Table 1: Summary of Dataset.

Dataset Summary			
Timespan	2007:2023	Authors	387
Sources	104	Author's Keywords	447
Documents	162	Single-authored documents	19
Avg citations per Document	57.19	Co-Authors per Document	3.09
References	11094	Document Average Age	4.97

disciplinary boundaries to establish itself as a globally pertinent research domain, rather than being confined solely within the purview of Information Systems (IS) research. Top of Form

The assessment of international collaboration is facilitated through the utilization of the Multi-Country Publication (MCP) ratio, which stands in contrast to the Single-Country Publication (SCP) metric, indicative of a lack of international partnership. The MCP ratio serves as a quantitative indicator of a nation's level of engagement in international collaboration, requiring the involvement of at least one co-author from a foreign country. Presented in Table 2 are MCP ratios for nations producing a minimum of two documents. The data showcased in Table 2 reveal that the United Kingdom, Italy, Norway, and France exhibit the highest levels of international collaboration suggesting a substantial presence of collaborative engagement among scholars in the field of BDAC research beyond national borders, denoting its significance beyond localized contexts. Conversely, nations such as the United States, China, India, and Germany demonstrate lower MCP ratios, potentially indicating a prevailing trend of reduced international collaboration within the IS scholarship of these respective countries.

Sources

As for the sources, 104 outlets published BDAC research during 2007-2023. Of the 104 sources, 74.2% were journals, and 21.4% were conference proceedings. We used Bradford's dispersion law to analyze the distribution of literature on a particular subject.[33] The law suggests that the sources in an academic field exhibit a known pattern (1: n: n^2) such that each of the three source clusters publishes about one-third of all the documents. The first cluster (or zone) is the core, followed by the second zone, with more sources, and the third zone covers the remaining and the bulk of the sources-each contributing to one-third of

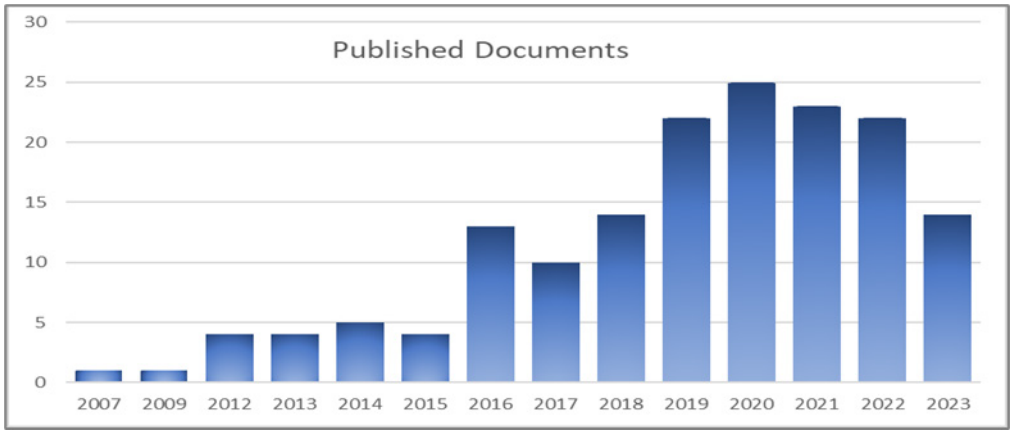

Figure 1: Number of Published Documents per year. (*Source: Prepared by the authors).*

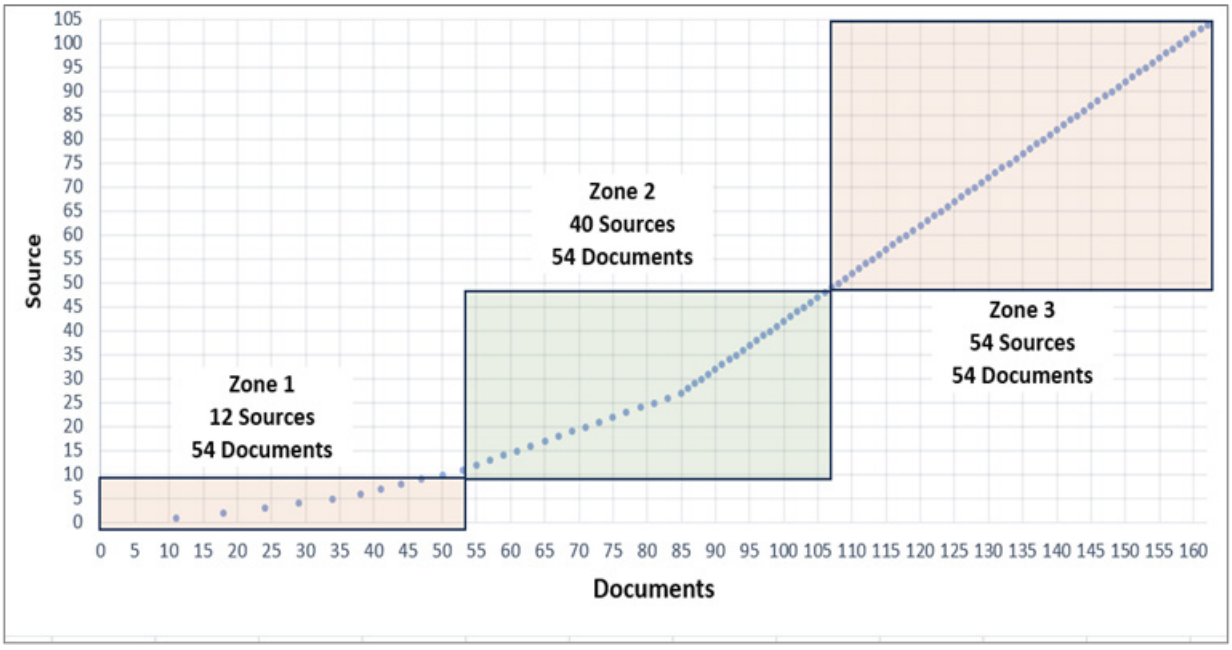

Figure 2: Bradford Dispersion Law Analysis.

documents. The core contains 12 distinct sources (~11.5%); the second zone includes 40 sources (~38.4%), and the third zone consists of 54 sources (~50%). Each zone accounts for one-third of the documents. The distribution, as illustrated in Figure 2, is approximately 1:3.3:4.3. Clearly, this distribution does not follow the Bradford law[33] of 1:3:9 (1: n: n^2), indicating that it is yet to reach the ideal concentration in terms of sources and documents. As the third zone articles are occasionally cited and offer marginal importance to the subject, it can be inferred that BDAC field is yet to achieve the law of diminishing returns.[33]

Authors

The Scopus data shows that 387 authors were responsible for the 162 documents, of which only 19 authors contributed to single-author documents. The top six authors by documents are Mikalef P.(10), Krogstie J.(8), Childe S J.(7), Dubey R.(8), Gunasekaran A.(7)

Table 2: Documents by Country and MCP Data.

Country	Documents	Freq	SCP	MCP	MCP Ratio
China	19	0.117	17	2	0.105
USA	17	0.105	11	6	0.353
United Kingdom	8	0.049	2	6	0.75
Australia	7	0.043	5	2	0.286
India	6	0.037	5	1	0.167
Jordan	6	0.037	5	1	0.167
Norway	6	0.037	0	6	1
United Arab Emirates	5	0.031	4	1	0.2
France	4	0.025	1	3	0.75
Colombia	3	0.019	2	1	0.333
Italy	3	0.019	1	2	0.667
Germany	2	0.012	2	0	0
Ghana	2	0.012	1	1	0.5
Iran	2	0.012	2	0	0
Ireland	2	0.012	1	1	0.5
Korea	2	0.012	1	1	0.5

* SCP-Single Country Publications, MCP-Multiple Country Publications.

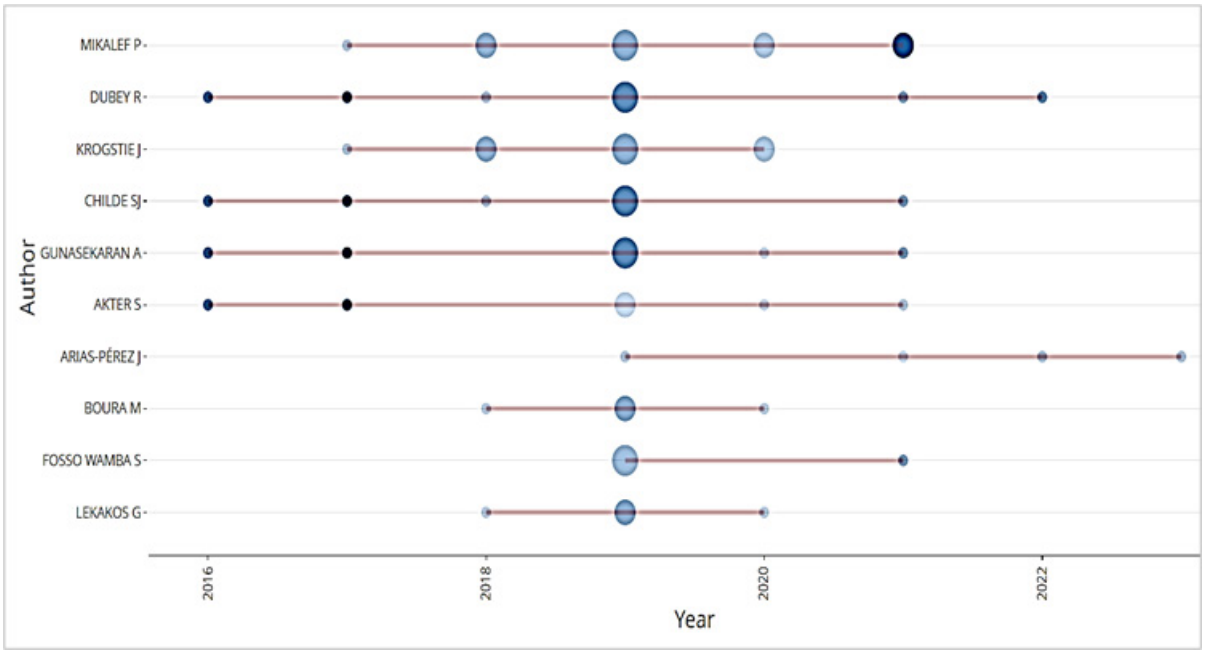

Figure 3: Scientific Productivity.

and Akter S.(6). Figure 3 shows the productivity trends of various authors and their timelines. It can be noted that the top 5 authors have collaborated on multiple documents. The intensity of the circle represents the total citations per year, and its size indicates the number of documents.

As observed from Figure 4, the top 5 cited documents from the citation analysis were Chen *et al.*,[2] Chen and Zhang,[35] Wamba *et al.*,[5] Wang *et al.*,[36] and Akter *et al.*[7] Evidently, these works shaped the research agenda of BDAC. While the first two papers were reviews and call for research on big data during the onset years, the later three papers were specific to BDAC phenomena. For example, Wamba *et al.*[5] and Akter *et al.*[7] use RBV, Information system (IS) success, and the IT business value research foundations to examine the effects of BDAC on firm performance. Other authors such as Wang and Hajli[37] and Wang *et al.*[36] focus on big data implementations in the healthcare domain. Unlike citation analysis, which focuses on the cited documents included in the research dataset, co-citation analysis relies upon the references cited by the documents. The co-citation analysis revealed the top 5 cited references, namely, Barney,[38] Teece *et al.*,[39] Bharadwaj,[40] Chen *et al.*,[2] and Bhatt and Grover.[41] They form the theoretical roots of the field (Figure 5). BDAC is grounded in strategic literature as indicated by RBV theory[38] and DCF,[39] IS resources, and capabilities literature.[40,41] Chen *et al.*,[2] was a call for research on the emerging phenomena of BDA.

Finally, we also validated the frequency distribution of scientific productivity using Lotka's[34] law, an inverse square law, which states that the number of authors making x contributions is about $100/x^p$ of those making one contribution, where p is ~2. Data shows that 329 authors contributed to only one article, while 58 authors contributed to two or more articles. In Figure 6, the dotted line represents the ideal authorship distribution as per Lotka's law, while the continuous line represents the distribution of authorship in BDAC. Evidently, BDAC authorship closely follows Lotka's law. The finding indicates the natural progression of the scholarship in line with the mature fields despite the novelty and addresses our first research question.

Analysis of Knowledge Structures

This step involves analyzing keywords, citation, and co-citation analysis using VOSviewer[28] to uncover the conceptual and intellectual structures. The co-occurrence of keywords analysis uses the keywords mentioned in the document to examine the conceptual structure of the research domain.[42] When words co-occur, it is assumed that the concepts are also closely related,

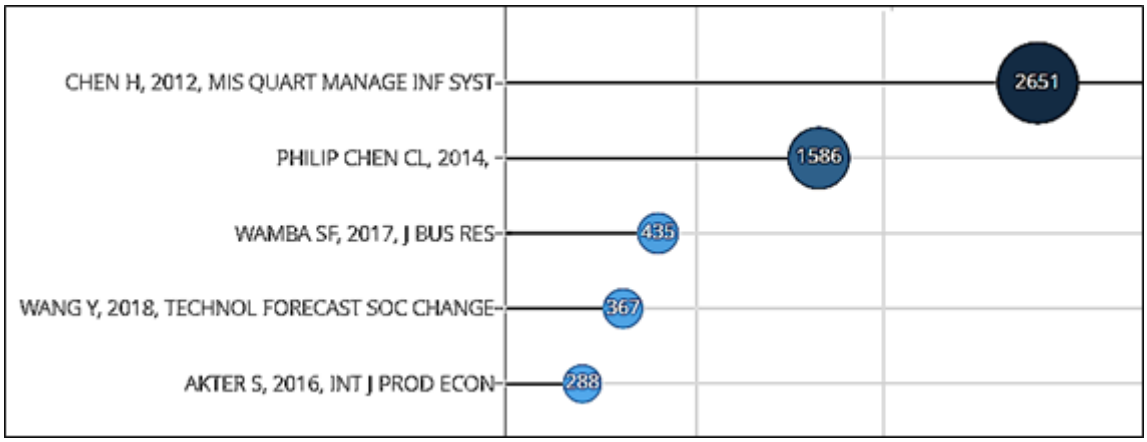

Figure 4: Most Cited Documents.

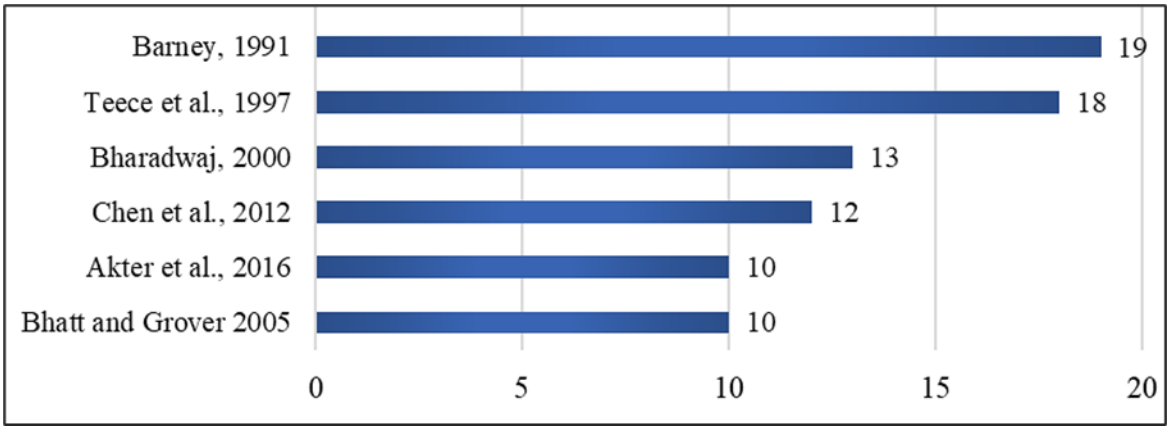

Figure 5: Most Cited References.

and this premise helps to uncover the conceptual structure and themes. The co-citation analysis measures the similarity using the frequency by which two documents are referenced by one or more documents.[43] When two documents quote the same reference, it suggests that the two documents address the same topic[44] and the greater the overlap, the stronger is the relationship between the two documents.

Documents

The citation and co-citation details of key documents are shown in Table 3. Data indicates that early works by Chen *et al.*,[2] Davenport,[45] and Chen and Zhang[35] shaped the research work in BDA. The multiple themes suggested by the authors in 2012 during the infancy of BDA helped establish a robust research agenda.

Further, the analysis of the references cited by the documents (via co-citation analysis) in Table 3 shows the theoretical pillars that have guided the scholarship in the last ten years. Among the 11094 references mentioned in Table 1, 45 references had at least five citations. The five most prominent references, which point to the vital theoretical logic of the field, are:

- Barney[38] on firm resources and competitive advantage.
- Teece *et al.*[39] on dynamic capabilities.
- Chen *et al.*[2] on business intelligence and analytics.
- Bharadwaj[40] on resource-based perspective on IS capability and firm performance.
- Bhatt and Grover[41] on types of IT capabilities.

It can be inferred that RBV and DCF are the two fundamental theories that are central to the examination of BDAC phenomena.

This is evident in the use of these theories in works by Gupta and George,[9] Wamba *et al.*,[5] Wang *et al.*[36] and Mikalef *et al.*[10] The citation documents network visualization with time overlay visualization (Figure 7) helps uncover the key trends and critical works that shaped the development of the field. The network graph indicates that the early works of Chen *et al.*[2] shaped the

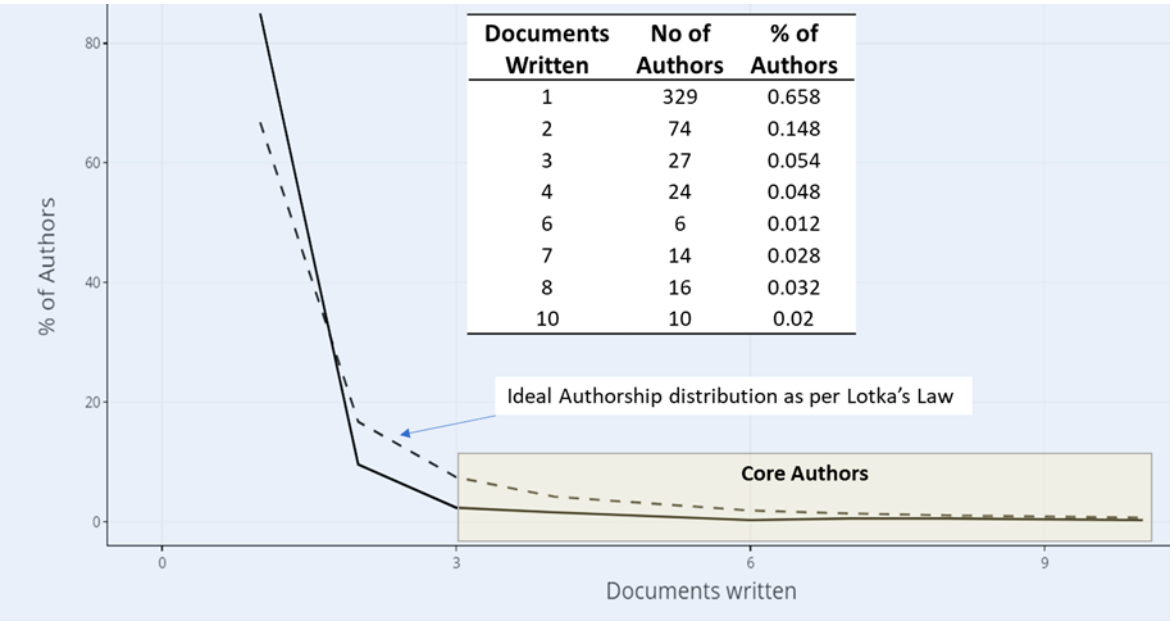

Figure 6: Frequency Distribution of Scientific Productivity.

Table 3: Analysis of Documents (Key Citations and Co-citations).

Citation Analysis		Co-Citation Analysis	
Document	**Total Citations**	**Document**	**Total Citations**
Chen H.[2]	2651	Barney, J[38]	19
Chen and Zhang[35]	1586	Teece, D. J.[39]	18
Wamba S. F.[5]	435	Bharadwaj A. S.[40]	13
Wang Y.[36]	367	Chen H.[2]	12
Akter S.[7]	288	Bhatt G. D.[41]	10
Gupta M.[9]	231	Akter S.[7]	10

research work in BDA. The next wave and progressive agenda were shaped by articles from Wamba *et al.*,[5] Wang *et al.*,[36] Akter *et al.*[29] and Gupta and George,[9] who provided further guidance and support in terms of operationalization and empirical research between 2015 and 2018. During this phase, scholarly attention gravitated towards the operationalization of BDAC constructs[9] and the examination of its direct impact on firm performance,[5,7] aligning with traditional Information Systems (IS) research themes. Nevertheless, it notably propelled the scholarly discourse forward, laying a robust groundwork for future investigations in the field.

Sources

The analysis of sources determines the journals that have contributed most to BDAC research. We analyze sources from the point of view of relevance (citation analysis) and foundations (co-citation analysis) (Table 4). Our dataset consisted of publications from 104 sources (journals and conferences) with the most significant number of citations from *Technological forecasting and social change (790), Journal of business research (692), International journal of production economics (657), Information and management (538)* with an average of 57.19 citations per doc. As far as co-citation is concerned (i.e., analyzing the cited journals by the documents in our data collection), studies have mostly cited documents from the *MIS Quarterly (357), Journal of Business Research (305), International Journal of Production Economics (261), Strategic Management Journal (248), and Journal of Management Information Systems (163).* The current publishing trends indicate that while the scientific investigation of BDAC is primarily published in journals related to the sources related to the IS management and strategic management literature such as *Strategic Management Journal* and *Harvard Business Review*, indicating the strategic importance of BDAC. In addition, the cross-functional nature of BDAC and its impacts on various organization functions also suggest BDAC publications in non-IS streams (such as operations and production economics).

Authors

As is expected in an emerging research field, authors have contributed substantially in the last few years. We find that 387 authors have contributed to 162 publications. Interestingly, the citation analysis reveals that the authors with the highest number of citations (2651) published only one article on the subject. A deeper analysis revealed that the MIS Quarterly article by Chen *et al.*[2] significantly impacted the field's growth by drawing research focus on BDA as a unique IS capability different from a traditional Business Intelligence capability. This is also evident in the growth of published documents (Figure 1).

The co-citation analysis in terms of authorship (i.e., authors cited in the reference list of the documents in the dataset) reveals that only 24 were cited more than 50 times. However, the data also shows a reliance on a few key scholars in IS and Strategy domain, such as Barney,[38] Teece *et al.*,[39] Bharadwaj,[40] Chen *et al.*,[2] Bhatt and Grover.[41] The Co-citation diagram density visualization (Figure 8) illustrates the clusters from left to the right showing the development of the field. The left cluster in red refers to the theoretical underpinnings of both IS and strategy management literature. The clusters on the right in yellow/ green color represent the recent developments in BDAC research such as the use of DCF[5,10] supply chain performance,[46] BDAC in healthcare.[36] This shift suggests the growing influence of BDAC, which is increasingly seen as a dynamic organizational capability and not just a typical IS capability.

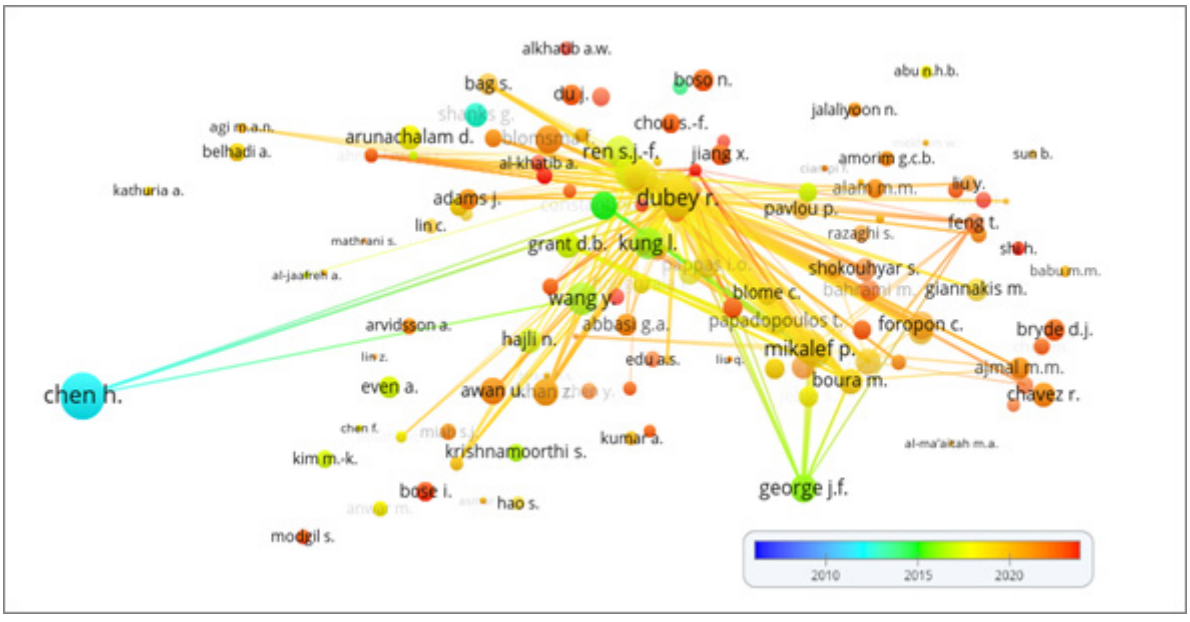

Figure 7: Citation Network with Time Overlay Visualization.

Table 4: Analysis of Top Sources.

Citation Analysis			Co-Citation Analysis	
Source	**# of Articles**	**Total Citations**	**Source**	**Total Citations**
Technological forecasting and social change.	11	790	MIS Quarterly	357
Journal of business research.	7	692	Journal of Business Research.	305
International journal of production economics.	5	657	International Journal of Production Economics.	261
Information and management.	6	538	Strategic Management Journal.	248
Sustainability (Switzerland).	5	73	Journal of Management Information Systems.	163
Business process management journal.	3	61	Harvard Business Review.	158
European journal of innovation management.	3	52	Journal of Operations Management.	120

Table 5: Analysis of Top Authors.

Citation Analysis		Co-Citation Analysis	
Author	**Total Citations**	**Author**	**Total Citations**
Chen *et al.*	2651	Gunasekaran A.	215
Chen P.	1586	Dubey R.	182
Wamba S. F.	435	Akter S.	145
Wang Y.	367	Childe S. J.	149
Akter S.	288	Wamba S. F.	158
Gupta M	231	Davenport R. H.	110
Isik O	177	Mikalef P.	102
Constantiou i.d.	176	Brynjolfsson E.	84
Awan u	161	Teece D. J.	79

In sum, most of the current research focuses on the value creation mechanisms of BDAC but used multiple theoretical logic to examine it. The evolving research themes indicate the emphasis on supply chain (resilience, agility, sustainability), manufacturing performance,[47] innovation, competitive performance, strategic value creation.[10,11] As for the theories, while RBV and DCF have been the prominent frameworks, researchers have approached BDAC from various theoretical perspectives as listed below:

- Absorptive Capacity (ACAP)[48]
- Affordance Theory[49]
- Contingency Theory and Structuration theory[50]
- Diffusion of Innovation Theory (DOI)[46]
- Entanglement View/Socio-materialism[5]
- Information Processing vie[51]
- Institutional Theory[47,52]
- Knowledge Management View (KM)[53,54]
- Social Capital Theory[54]
- Socio-Technical System Theory (STS)[49] and
- Technology, Organization, and Environment (TOE).[55]

A key takeaway from our analysis is the increasing reference to DCF (Table 3) which is evident in recent studies. These works have focused on indirect effects such as agility, knowledge management, and innovation.[5,10] In another interesting paper by Delen and Zolbanin,[56] attention is drawn towards the exploration and incorporation of BDA as a complementary research paradigm within the domain of business studies. While contemporary business challenges often elude resolution through conventional research methodologies, skepticism persists in certain academic domains regarding the efficacy of analytical methodologies in addressing such issues.

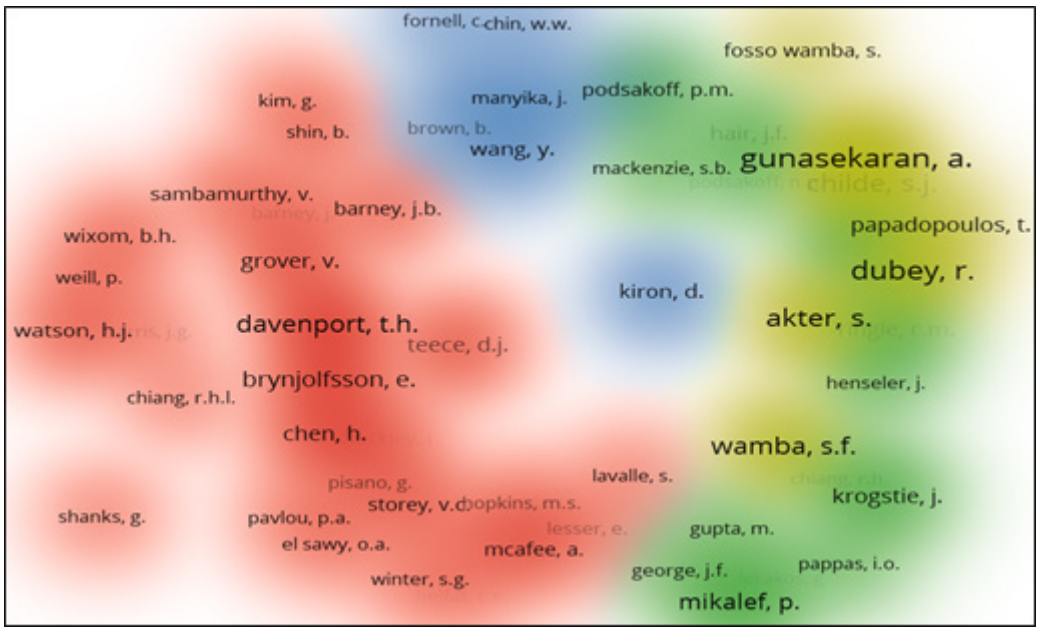

Figure 8: Co-citation Authors-Heatmap/Density Visualization.

Keywords

The keyword growth of the field in the last ten years is illustrated in the time overlay graph (Figure 9). For our analysis, in VOSviewer, similar-sounding terms were deleted or merged using a dictionary file. For example, words such as "structural equation modeling" and "least squares approximations" and duplicate phrases such as "Big Data Analytics" and "Data Analytics" were removed. The time overlay graph (Figure 9) shows keyword growth with time and indicates the importance of DCF compared to RBV, which has been the key theory in many IS studies. This finding corroborates with the high co-citations (Table 3) of the works by Teece *et al.*[39] and Barney.[38] Further, in line with DCF studies, the focus shifted from firm performance, information management, decision making to new emerging themes, such as innovation, agility, knowledge management under the business value umbrella[5,10,57] These findings are consistent with the rationale underpinning DCF and underscore the significance of BDAC as an organizational capability rather than solely an IS capability. Also, it can be inferred that BDAC is not just an Information systems capability but a critical organization capability due to its strategic potential. Consequently, the bibliometric analysis addresses our second research question regarding the evolution of BDAC research.

The exploration of the keyword co-occurrence network via VOSViewer exposes eight distinct clusters (Figure 10). The smaller clusters, comprising fewer than three elements, were merged with larger clusters, resulting in five clusters (Table 5). The clusters represent five research themes discussed below:

Theme 1 (BDAC Antecedents and Consequences)

This theme represents studies on BDAC effects, namely business value, competitive advantage, firm performance, BDAC antecedents and factors that impact this relationship organization culture, Strategic alignment, RBV theory.[5,7] Additionally, studies have also focused on BDAC constituents such as human, technical, process[9] and IT-business process integration.

Theme 2 (Business Intelligence)

This theme reflects on the traditional view of business analytics, namely business intelligence, and its relationship with decision making, information processing and management, knowledge management, and firm performance.[2,50]

Theme 3 (BDAC and Environment Dynamics)

This theme focuses on the environment dynamics and its linkages with BDAC. In line with the antecedents of DCF, this cluster has concentrated on investigating the indirect effects of BDAC, such as ambidexterity, agility, innovation, and environmental orientation.[5,10]

Theme 4 (Advanced Analytics)

This theme represents studies on advanced analytics such as AI, predictive analytics, social data and analytics, new data architectures in domains such as a smart city. The studies in this cluster primarily focusses on data, tools, and technologies such as AI, ML, and social data and their effects on business performance.

Theme 5 (BDAC Process and Industry Focus)

This theme reflects the effects and use of BDAC in specific business processes such as supply chain[52] and healthcare.[36] The cluster emphasizes BDAC as critical to various business functions and processes at the industry level.

The five research themes emerging from these clusters answers our third research question.

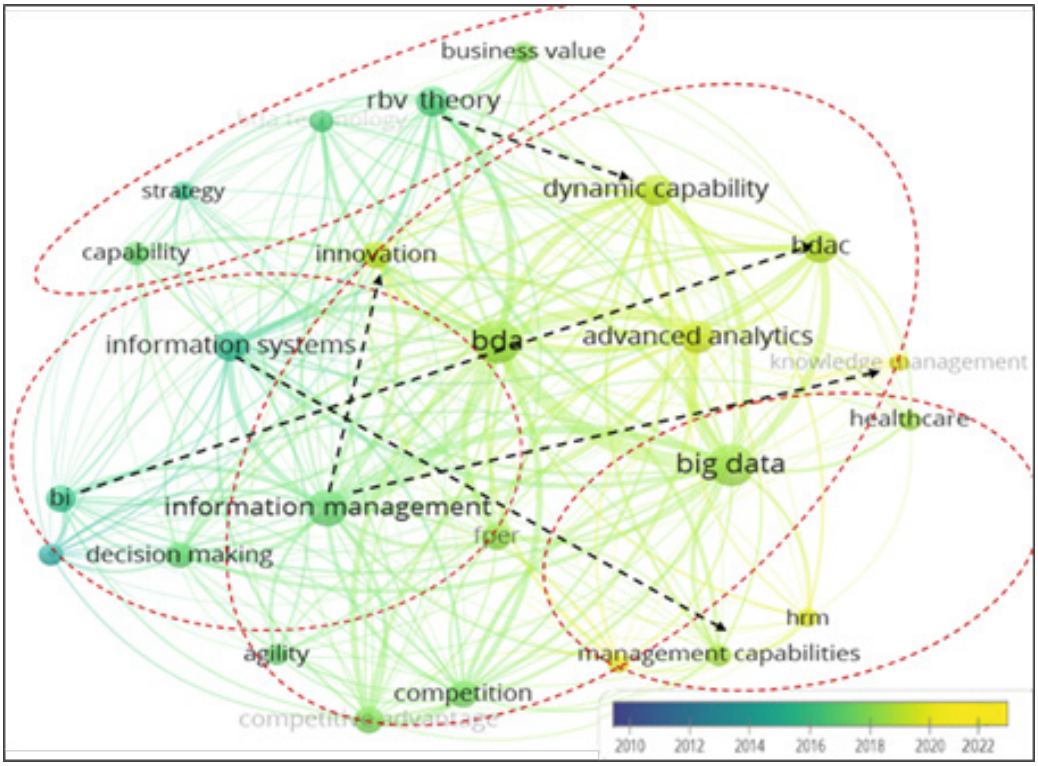

Figure 9: Keyword Analysis.

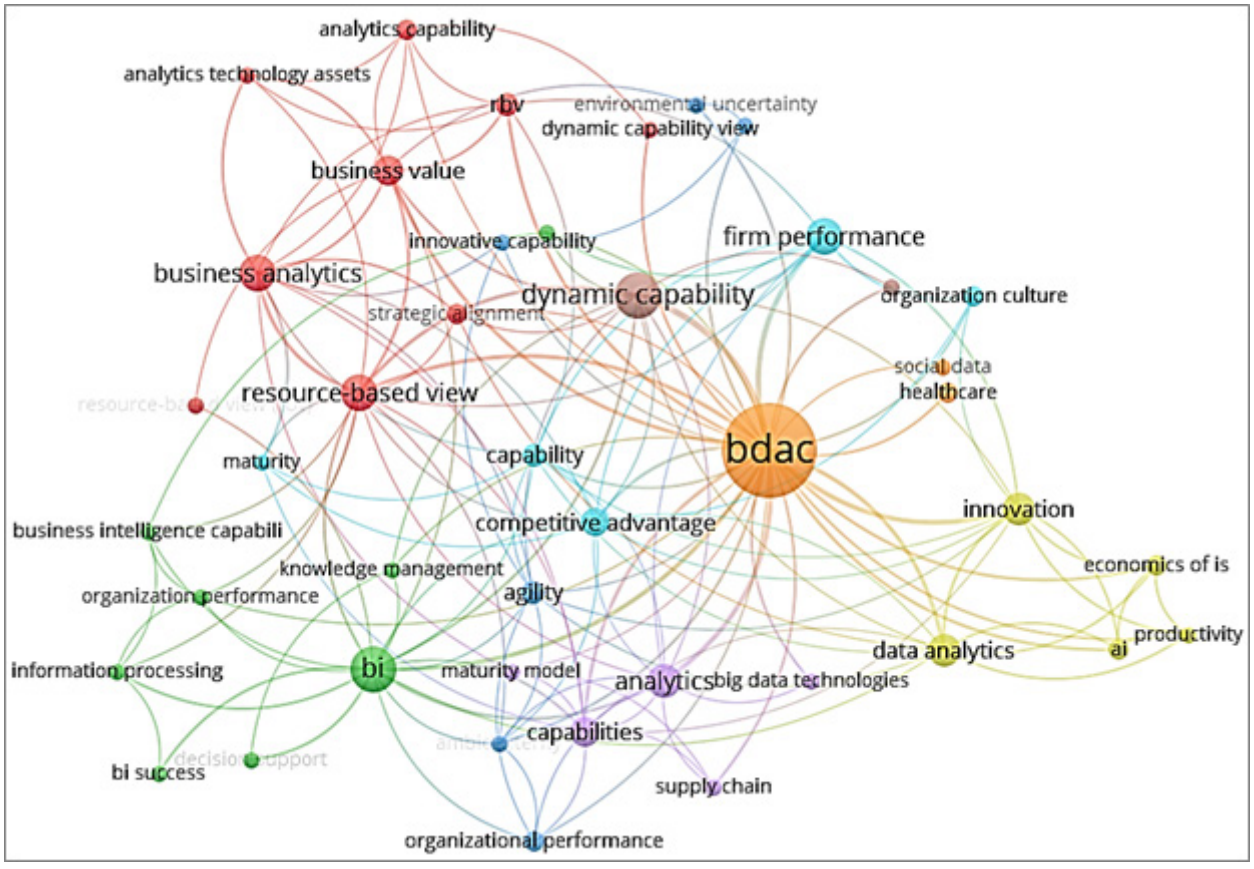

Figure 10: Keyword Co-occurrence Network Diagram.

Table 5: Keyword Analysis-Clusters.

C1: BDAC Antecedents and Impacts	C3: BDAC and Environment Dynamics
Analytics Capability	Agility
Analytics Technology Assets	Ambidexterity
Business Analytics	Dynamic Capability
Business Value	Environment Orientation
Competitive Advantage	Environmental Uncertainty
Firm Performance	Innovative Capability
Maturity	Information Governance
Organization Culture	C4: Advanced Analytics
Resource Based View	Artificial Intelligence
Strategic Alignment	Data Analytics
C2: Business Intelligence	Innovation
Business intelligence capability	Economics of IS
BI and Business Value	Productivity
BI success	C5: Process and Industry Impacts
Decision Support	Analytics
Information Processing	Big Data Technologies
Knowledge Management	Capabilities
Maturity Model	Healthcare
Organizational Performance	Supply Chain
Organizational Agility	Social Data

Below, we summarize the overall findings from the analysis, followed by implications, and research opportunities in the next section.

BDAC is an emerging research agenda indicated by growth in the last few years with a global research footprint and international collaboration as indicated by the MCP ratio.

The 23% share of business/management research articles indicates a focus on non-technical factors and evolution as a multidisciplinary research field.

Scientific productivity was found to be in line with Lotka's law showing signs of maturity despite the novelty.

The strategic management literature influences the theoretical foundations of BDAC. Due to BDA's changing nature, BDAC issues have been approached from different theoretical perspectives.

The two prominent and widely used theories used in BDAC studies emerging from our review are RBV[38] and DCF.[39] The research objectives in existing studies were found to be in line with the underlying theory used. While RBV based studies focus on the direct effects of BDAC such as firm performance, information management, new emerging themes focus primarily on indirect effects such as innovation, agility, and knowledge management.

IMPLICATIONS AND FUTURE RESEARCH OPPORTUNITIES

Implications for Research

BDAC Conceptualization

Given the nascent stage of BDAC research, there is an opportunity to contribute to the advancement of the field by developing new research frameworks, models, and measurement instruments tailored to its unique characteristics and challenges. A focus on the micro-foundation capabilities of BDAC in line with the DCF logic can have significant implications for BDAC research. Further, consideration of BDAC as a DC (Theme 3) can offer new and diverse perspectives on BDAC business value research, facilitating the examination of indirect business value creation (Theme 1, 4).

BDAC Antecedents

Our finding suggests that there is a need for a broader investigation of the potential ways big data can create business value (Theme 1, 4). Examining business value from diverse perspectives of big data utilization, driving process improvement, a strategic tool for creating new data-based products or services, and creation of indirect business value such as innovation, new data centric business models have far-reaching implications for BDAC scholarship and industry. These aspects have received limited attention in existing research.

Industry and Firm

At an industry and firm level, an exploration of the ex-ante character of BDAC in various organizational processes, including strategy formulation, decision-making, capability development, and adoption (Theme 1), can shed light on BDAC development, challenges and opportunities for successful BDAC adoption. These aspects are critical to effectively leverage BDAC for organizational effectiveness and gaining competitive advantage.

Further, existing research predominantly emphasizes on the mere existence of BDAC and direct value creation measures such as firm performance. Assessing the impact of firm-specific actions, such as organic versus inorganic growth strategies (Theme 1) can enhance the understanding of the factors influencing BDAC evolution for strategic planning and resource allocation.

Finally, as BDAC matures in a firm, an examination of BDAC maturity (Theme 2) and its effects on various business processes across different industry contexts can provide insights into the firm and industry-specific factors critical for ensuring BDAC success, guiding managerial decisions and informing organizational strategies.Top of Form

Managerial Implications

Despite the purported benefits of BDAC, many firms fail to capture value from BDA initiatives,[6] suggesting investigation of critical success factors of BDAC and why some firms are better at extracting business value from BDAC in comparison to others. Our study yields some interesting implications valuable for the practitioners and policymakers engaged in BDA.

First, from the viewpoint of capabilities, BDA should not be seen as just another technology capability but a key strategic capability that can help in a) driving efficiencies and b) as a strategic tool to create new data-based products and new business models. The efficiency benefits are keys to implement a low-cost strategy, while the new data-based products are critical for a growth strategy. Thus, managers and leaders should take a long term multi-dimensional strategic view in planning, and investing in big data initiatives.

Second, from a business value standpoint, the BDA value can manifest within and outside the firm in tangible and intangible forms. Therefore, managers should include both tangible and intangible benefits while making BDAC investment decisions, business case justification and evaluating the success (or failure) of BDAC projects.

Finally, our review suggests that adequate consideration should be given to non-technological factors (Cluster C1) and industry-specific factors (Cluster C5) to develop BDA capability. Therefore, managers should focus on people, skills, technology, process, and industry/firm-specific factors to define a tailored approach for BDAC implementation in their firm.

Future Research Directions

Based on the analysis, identified themes and implications, we present some interesting avenues for future research:

Investigating the conceptualization of BDAC as a DC and its subsequent indirect impacts on diverse business value dimensions and its strategic implications presents numerous promising avenues for future research. A shift from the traditional RBV approach to a DCF based framework for understanding BDAC could prove to be a compelling research endeavor aimed at exploring the strategic significance of BDAC.

Extending this line of enquiry, scholars can also explore the development of novel measurement instruments (based on DCF) and methodologies for assessing the effectiveness and impact of BDAC, providing researchers and practitioners with valuable insights and interesting research areas. Researchers can explore and undertake empirical and cross-sectional studies focusing on various micro-foundations of BDAC DC and its effects.

Given that 80% of BDA initiatives will fail to deliver any business benefits,[6] a fascinating area of research is to analyze and shed light on the factors contributing to unsuccessful BDA initiatives and suggesting strategies for mitigating risks and enhancing project success rates.

Continuing this line of inquiry, scholars could delve into the maturity of BDAC and its correlation within the industry and organizational factors, contributing to differential levels of BDAC success among firms.

As BDAC matures within a firm, a scholarly investigation into how BDAC facilitates the reconfiguration of other organizational capabilities to adeptly respond to dynamic market conditions and competitive pressures could also provide valuable practical insights from a strategic perspective.

Finally, researchers can investigate diverse ways in which big data can be leveraged to create new business models, new value streams, process innovation across various industries and organizational contexts.

CONCLUSION AND LIMITATIONS

The objective of this work was to examine and provide a comprehensive view of the BDAC research, trace the evolution of the field, and identify key research trends and opportunities We used Bibliometrix R and VOSviewer software to analyze the descriptive features of the dataset using technique such as Lotka's law Bradford dispersion law and Multi-Country Publication (MCP) ratio revealing maturity and productivity characteristics. The analysis of various documents, keywords, authors, and sources using citation, co-citation, and keyword analysis helped us to understand the key research trends and evolution of BDAC research over time. Results indicate BDAC research's emergence and evolution over the past decade, aligning with Lotka's law, but not Bradford's dispersion law, suggesting a maturing field with growth potential. International collaboration, assessed through the multi-country publication ratio, highlights significant global engagement, particularly in the UK, Australia, Norway, and France. Analyzing keyword growth over time revealed the roots and emerging trends in BDAC research. Traditionally, studies focused on the business value of big data, business intelligence, firm performance, and information management, treating BDAC as a conventional IS capability. Recent research, however, diverges by considering BDAC as an organizational capability. A notable trend is the increasing use of DCF to explore the indirect effects of BDAC, including innovation and agility.

The theoretical contributions of the study pertain to the consolidation of existing works. It identifies five BDAC research clusters: a) BDAC antecedents and consequences, b) business intelligence, c) BDAC and environment dynamics, d) advanced analytics, and e) BDAC process and industry focus. The key trends, knowledge structures, theoretical roots and thematic clusters offer new research directions which include but not limited to BDAC positioning as a DC, DC based BDAC conceptualization, micro foundational capabilities of BDAC, new ways oof BDAC

operationalization, BDAC response to market dynamics with a focus on BDAC successes and failures, BDAC maturity, BDAC critical success factors, and BDAC effects on other capabilities for a strategic roleplay. A limitation of our study is reliance on limited qualitative interpretation. Future work could delve into detailed qualitative studies on each theme.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

1. Bharadwaj A, El Sawy OA, Pavlou PA, Venkatraman N. Digital business strategy: toward a next generation of insights. MIS Q. 2013;37(2):471-82. doi: 10.25300/MISQ/2013/37: 2.3.
2. Chen H, Chiang RH, Storey VC. Business intelligence and analytics: from big data to big impact. MIS Q. 2012;36(4):1165-188. doi: 10.2307/41703503.
3. Grover V, Chiang RH, Liang TP, Zhang D. Creating strategic business value from big data analytics: A research framework. J Manag Inf Syst. 2018;35(2):388-423. doi: 10.1080/07421222.2018.1451951.
4. Ransbotham S. Analytics as a source of business innovation [Internet]; 2017. MIT Sloan Management Review [cited May 16 2019]. Available from: https://sloanreview.mit.edu/projects/analytics-as-a-source-of-business-innovation/.
5. Wamba SF, Gunasekaran A, Akter S, Ren SJ, Dubey R, Childe SJ. Big data analytics and firm performance: effects of dynamic capabilities. J Bus Res. 2017;70:356-65. doi: 10.1016/j.jbusres.2016.08.009.
6. White A. Our Top Data and analytics predicts for 2019 [Internet]. Gartner. Gartner; 2019. Available from: https://blogs.gartner.com/andrew_white/2019/01/03/our-top-data-and-analytics-predicts-for-2019/.
7. Akter S, Wamba SF, Gunasekaran A, Dubey R, Childe SJ. How to improve firm performance using big data analytics capability and business strategy alignment? Int J Prod Econ. 2016;182(1):113-31. doi: 10.1016/j.ijpe.2016.08.018.
8. Khanra S, Dhir A, Mäntymäki M. Big data analytics and enterprises: a bibliometric synthesis of the literature. Enterpr Inf Syst. 2020;14(6):737-68. doi: 10.1080/17517575.2020.1734241.
9. Gupta M, George JF. Toward the development of a big data analytics capability. Information &and Management. 2016;53(8):1049-64.
10. Mikalef P, Boura M, Lekakos G, Krogstie J. Big data analytics and firm performance: findings from a mixed-method approach. J Bus Res. 2019;98:261-76. doi: 10.1016/j.jbusres.2019.01.044.
11. Mikalef P, Pappas IO, Krogstie J, Giannakos M. Big data analytics capabilities: a systematic literature review and research agenda. Inf Syst e-Business Manag. 2018;16(3):547-78. doi: 10.1007/s10257-017-0362-y.
12. Torres R, Sidorova A, Jones MC. Enabling firm performance through business intelligence and analytics: A dynamic capabilities perspective. Information & Management. 2018;55(7):822-39. doi: 10.1016/j.im.2018.03.010.
13. Phillips-Wren G, Iyer LS, Kulkarni U, Ariyachandra T. Business analytics in the context of big data: A roadmap for research. Commun Assoc Inf Syst. 2015;37(23):448-72. doi: 10.17705/1CAIS.03723.
14. Singh N14. Singh N. Big data technology: developments in current research and emerging landscape. Enterpr Inf Syst. 2019;13(6):801-31. doi: 10.1080/17517575.2019.1612098.
15. Qazi RU, Sher A15. Sher A. Big Data applications in businesses: an overview. Int Technol Manag Rev. 2016;6(2):50. doi: 10.2991/itmr.2016.6.2.3.
16. Chong D, Shi H16. Shi H. Big data analytics: a literature review. J Manag Anal. 2015;2(3):175-201. doi: 10.1080/23270012.2015.1082449.
17. Sheng J, Amankwah-Amoah J, Wang X. A multidisciplinary perspective of big data in management research. Int J Prod Econ. 2017;191(1):97-112. doi: 10.1016/j.ijpe.2017.06.006.
18. Pathak S, Krishnaswamy V, Sharma M. Big data analytics capabilities: a novel integrated fitness framework based on a tool-based content analysis. Enterpr Inf Syst. 2023;17(1):1-35. doi: 10.1080/17517575.2021.1939427.
19. Paul J, Merchant A, Dwivedi YK, Rose G. Writing an impactful review article: what do we know and what do we need to know? J Bus Res. 2021;133(3):337-40. doi: 10.1016/j.jbusres.2021.05.005.
20. Karimi J, Walter Z. The role of dynamic capabilities in responding to digital disruption: A factor-based study of the newspaper industry. J Manag Inf Syst. 2015;32(1):39-81. doi: 10.1080/07421222.2015.1029380.
21. Vial G. Understanding digital transformation: a review and a research agenda. J Strateg Inf Syst. 2019;28(2):118-44. doi: 10.1016/j.jsis.2019.01.003.
22. Redman TC. data driven: profiting from your most important business asset. Boston: Harvard Business Press. Cop; 2008.
23. Lycett M. 'Datafication': Making sense of (big) data in a complex world. European Journal of Information Systems. 2013;22(4):381-6.
24. Kwon O, Lee N, Shin B. Data quality management, data usage experience and acquisition intention of big data analytics. Int J Inf Manag. 2014;34(3):387-94. doi: 10.1016/j.ijinfomgt.2014.02.002.
25. Ghasemaghaei M, Hassanein K, Turel O. Impacts of big data analytics on organizations: A resource fit perspective; 2015. In: AMCIS 2015 proceedings [Internet]. Available from: https://aisel.aisnet.org/amcis2015/BizAnalytics/GeneralPresentations/19.
26. Kiron D. The analytics mandate [Internet]; 2014. MIT Sloan Management Review [cited Dec 27 2019]. Available from: https://sloanreview.mit.edu/projects/analytics-mandate/.
27. Mongeon P, Paul-Hus A. The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics. 2016;106:213-28.
28. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 2010;84(2):523-38. doi: 10.1007/s11192-009-0146-3, PMID 20585380.
29. Aria M, Cuccurullo C29. Bibliometrix: an R-tool for comprehensive science mapping analysis. J Inf. 2017;11(4):959-75. doi: 10.1016/j.joi.2017.08.007.
30. Rogers G, Szomszor M, Adams J. Sample size in bibliometric analysis. Scientometrics. 2020;125(1):777-94. doi: 10.1007/s11192-020-03647-7.
31. Abramo G, D'Angelo CA, Viel F. The field-standardized average impact of national research systems compared to world average: the case of Italy. Scientometrics. 2011;88(2):599-615. doi: 10.1007/s11192-011-0406-x.
32. Mayr P, Scharnhorst A. Scientometrics and information retrieval: weak-links revitalized. Scientometrics. 2015;102(3):2193-9. doi: 10.1007/s11192-014-1484-3.
33. Bradford SC. Classic paper: Sources of information on specific subjects. Collection Management. 1976;1(3-4):95-104.
34. Lotka AJ. The frequency distribution of scientific productivity. J Franklin Inst. 1926 Aug;202(2):271.
35. Chen CL, Zhang CY. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Inf Sci. 2014;275:314-47. doi: 10.1016/j.ins.2014.01.015.
36. Wang Y, Kung L, Byrd TA. Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. Technol Forecasting Soc Change. 2018;126:3-13. doi: 10.1016/j.techfore.2015.12.019.
37. Wang Y, Hajli N. Exploring the path to big data analytics success in healthcare. J Bus Res. 2017;70:287-99. doi: 10.1016/j.jbusres.2016.08.002.
38. Barney J38. Barney J. Firm resources and sustained competitive advantage. J Manag. 1991;17(1):99-120. doi: 10.1177/014920639101700108.
39. Teece DJ, Pisano G, Shuen A. Dynamic capabilities and strategic management. Strateg Manag J. 1997;18(7):509-33. doi: 10.1002/(SICI)1097-0266(199708)18:7<509:AID-SMJ882>3.0.CO;2-Z.
40. Bharadwaj AS. A resource-based perspective on information technology capability and firm performance: an empirical investigation. MIS Q. 2000;24(1):169. doi: 10.2307/3250983.
41. Bhatt GD, Grover V. Types of information technology capabilities and their role in competitive advantage: an empirical study. J Manag Inf Syst. 2005;22(2):253-77. doi: 10.1080/07421222.2005.11045844.
42. Callon M, Courtial JP, Turner WA, Bauin S. From translations to problematic networks: an introduction to co-word analysis. Soc Sci Inf. 1983;22(2):191-235. doi: 10.1177/053901883022002003.
43. McCain KW. Mapping authors in intellectual space. Commun Res. 1989;16(5):667-81. doi: 10.1177/009365089016005007.
44. Kessler MM. Bibliographic coupling between scientific papers. American Documentation. 1963;14(1):10-25.
45. Davenport TH, Barth P, Bean R. How Big Data Is Different. MIT Sloan Management Review. 2012;54(1):43. Available from: https://sloanreview.mit.edu/article/how-big-data-is-different/.
46. Gunasekaran A, Papadopoulos T, Dubey R, Wamba SF, Childe SJ, Hazen B, *et al.* Big data and predictive analytics for supply chain and organizational performance. J Bus Res. 2017;70(1):308-17. doi: 10.1016/j.jbusres.2016.08.004.
47. Dubey R, Gunasekaran A, Childe SJ. Big data analytics capability in supply chain agility. Manag Decis. 2019;57(8):2092-112. doi: 10.1108/MD-01-2018-0119.
48. Božič K, Dimovski V48. Business intelligence and analytics for value creation: the role of absorptive capacity. Int J Inf Manag. 2019;46:93-103. doi: 10.1016/j.ijinfomgt.2018.11.020.
49. Dremel C, Herterich MM, Wulf J, vom Brocke J. Actualizing big data analytics affordances: A revelatory case study. Information &and Management. 2018;57(1):103121.
50. Kulkarni U, Robles-Flores J, Popovič A. Business intelligence capability: the effect of top management and the mediating roles of user participation and analytical decision making orientation. J Assoc Inf Syst. 2017;18(7):516-41. doi: 10.17705/1jais.00462.
51. Cao G, Duan Y, Li G. Linking business analytics to decision making effectiveness: A path model analysis. IEEE Trans Eng Manage. 2015;62(3):384-95. doi: 10.1109/TEM.2015.2441875.
52. Dubey R, Gunasekaran A, Childe SJ, Blome C, Papadopoulos T. Big data and predictive analytics and manufacturing performance: integrating institutional theory,

resource-based view and big data culture. Br J Manag. 2019;30(2):341-61. doi: 10.1111/1467-8551.12355.

53. Côrte-Real N, Oliveira T, Ruivo P. Assessing business value of Big Data Analytics in European firms. J Bus Res. 2017;70:379-90. doi: 10.1016/j.jbusres.2016.08.011.
54. Malgonde O, Bhattacherjee A. Innovating using big data: A social capital perspective research-in-progress [Internet]; 2014. Available from: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1672&andcontext=amcis2014.
55. Chen DQ, Preston DS, Swink M. How the use of big data analytics affects value creation in supply chain management. J Manag Inf Syst. 2015;32(4):4-39. doi: 10.1080/07421222.2015.1138364.
56. Delen D, Zolbanin HM. The analytics paradigm in business research. J Bus Res. 2018;90:186-95. doi: 10.1016/j.jbusres.2018.05.013.
57. Mikalef BM, Lekakos G, Krogstie J. The Role of Information Governance in Big Data Analytics driven Innovation. Information &and Management. 2020;39:103361.

Cite this article: Pathak S, Krishnaswamy V, Sharma M. Unveiling the Evolution of Big Data Analytics Capability: A Comprehensive Analysis. J Scientometric Res. 2024;13(3):791-805.