

## Article

# The Nexus between Big Data and Sustainability: An Analysis of Current Trends and Developments

Fernando Garrigós-Simón <sup>1</sup>, Silvia Sanz-Blas <sup>2,\*</sup>, Yeamduan Narangajavana <sup>3</sup> and Daniela Buzova <sup>2</sup>

<sup>1</sup> Departamento de Organización de Empresas, Universitat Politècnica de València, 46022 Valencia, Spain; fgarrigos@doe.upv.es

<sup>2</sup> Departamento de Comercialización e Investigación de Mercados, Universitat de Valencia, 46022 Valencia, Spain; daniela.buzova@uv.es

<sup>3</sup> Departamento Administración de Empresas, Universitat Jaume I, 12071 Castellón, Spain; ynaranga@uji.es

\* Correspondence: silvia.sanz@uv.es

**Abstract:** With the development of technological innovations, Big Data is transforming the socio-economic world, impacting almost every organization and person. The transformations associated with the development of Big Data have important consequences for the sustainability of organizations, regions, and the society as a whole, and as such, they have been specifically addressed by the academic literature focusing on sustainability. Despite its importance, and perhaps because of its rapid emergence, there is a lack of studies dealing with the analysis of this body of literature and its trends. The current research attempts to fill this gap. The study develops a bibliometric and visualization analysis of the literature on the nexus between Big Data and Sustainability. The research analyzes 726 documents on this topic, published until the end of 2020, in the Web of Science Core Collection database through the VOSviewer software. The results indicate the main trends and developments on the topic related to the most cited papers, authors, publications, institutions, and countries. The visualized frameworks, structures and trends are useful for both researchers and practitioners, as they can help them understand the current situation, issues to consider, and main developments on the topic.

**Keywords:** big data; sustainability; bibliometric; visualization; trends



**Citation:** Garrigós-Simón, F.; Sanz-Blas, S.; Narangajavana, Y.; Buzova, D. The Nexus between Big Data and Sustainability: An Analysis of Current Trends and Developments. *Sustainability* **2021**, *13*, 6632. <https://doi.org/10.3390/su13126632>

Academic Editor: Roberto Cerchione

Received: 11 May 2021  
Accepted: 3 June 2021  
Published: 10 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Big Data is very important for organizations and people, as its developments and associated transformations are not only enhancing long term innovation, competitiveness, and sustainability of organizations and regions but also changing the functioning and the understanding of nowadays society. In the new Ubiquitous Environment, derived from the “The Internet of Things”, “Pervasive and Ubiquitous Computing”, and the evolution of the semantic or “Web 3.0 Era” [1], the development of “artificial intelligence” and new robots, and technological improvements, people, scholars, and organizations need to understand these transformations to be able to adapt to them.

However, the search for sustainable big data developments, conceived as Big Data developments that follow a sustainable perspective, requires a deep understanding of the meaning of both terms (concepts and evolutions), the existing perspectives of their conceptualizations, and stakeholders’ perceptions. Specifically, the diverse sustainability conceptualizations have to be considered in order to observe the sustainability of these developments, or the conception of sustainable development goals. In addition, the evolution of the technology, and the processes related to big data, must be also observed. This is essential because the economic and financial sustainability of organizations and areas, as well as the natural, human, and social capital sustainability [2] also depend on this knowledge because innovative Big Data applications can also be an important mechanism to enhance sustainability.

On the basis of the above, the aim of this research is to develop a deep analysis of the conceptualizations, state of the art and the emerging trends in the literature on the nexus between big data and sustainability. The paper also attempts to structure, and graphically illustrate, the research that relates big data and sustainability, through the development of a bibliographic and visualization analysis. In this way, both scholars and practitioners can understand the complexities and interactions when dealing with these issues.

To achieve the aforementioned objectives, this section briefly explains the importance of developing a bibliometric study to analyze BD&S research. After emphasizing the lack of previous review and/or bibliometric analyses of both terms, the section highlights the importance and contributions of our work. The next section introduces the main concepts and perspectives associated with big data and sustainability, pointing out the relevance of big data in sustainability research. Once the methodology has been described, the following sections develop the proposed bibliometric and visualization analysis.

Although there is an exponential growth in both the big data and the sustainability literature, little is known about the current state of knowledge on the relationship between the two fields of research. Bibliometric analyses are widely recognized and extended methods in the literature for such purposes. Specifically, these analyses are conceived as a cross-disciplinary science focused on quantitative analysis of bibliographic data through the use of statistical and mathematical tools [3]. The academy highlights the ability of this technique to analyze specific research areas from objective information in an easily treatable way [2]. The technique has been broadly used to identify and describe the state of the art and trends of a wide range of topics and disciplines, as well as networks and connections among topics, authors, institutions, countries, and research areas. Hence, we consider bibliometric analysis as an appropriate method to analyze and evaluate the quality and the research interest in the intersection between the two topics.

There are several bibliometric studies that identify the progress of Big Data related issues in general [4–10], Big Data technologies [11], or Big Data analytics [12–14]. There are also analyses on the development of big data in different fields, such as “Medical Big Data research” [15], “Big data analytics in health” [16], “Big Data research in healthcare informatics” [17], “Map Reduce in the field of Big Data” [18], “Big Data and Machine learning” [19], “Big Data and enterprises” [20], “Big Data sources and control charts” [21], “Big Data and Business Processes” [22], “Big Data and supply chain management” [23,24], “Big Data intelligence” [25], “Business Intelligence and Analytics” [26,27], “Big Data and Analytics, and performance” [28], “Big Data & Dynamic capabilities” [29], “Big Data and consumer behavior” [30].

We have also found bibliometric studies connecting Big Data with specific research areas such as Bioinformatics [31], Blockchain research [32], Internet of things research [33], open data [34], Industry 4.0 Research [35–37], industry 4.0 technologies [38], Artificial intelligence [39], and Data science [40]. There have also been bibliometric analyses that look at the relationship between big data and certain sustainability-related aspects such as: the impact of big data and internet of things on the circular economy [41], Industry 4.0, process safety and environmental protection [42], sustainable industry 4.0 [43], open innovation and sustainable tourism [44], sustainable manufacturing and industry 4.0 [45], smart cities [14,46], urban sustainability [47], and sustainable supply chain [48,49]. However, we could not find any bibliometric or visualization study focusing on big data and sustainability.

Despite the current lack of BD&S research, understanding the relationship between Big Data and Sustainability is crucial. It is essential for authors and researchers to understand the state of the art of this literature, the main and new trends that connect Big Data and Sustainability, as well as observe the main perspectives and streams of research that have analyzed the connection of both terms, the pioneering studies, the topics that are underdeveloped, and the impacts and relationships among existing studies. In this regard, the present work can uncover the structure of these analyses and their developments, which is key in shaping and planning future research. In addition, the study is also relevant for practitioners and policy makers since the development of new business strategies, and

public policies, related to Big Data must incorporate sustainability aspects. This, in turn, is vital for enhancing not only the company image and competitiveness, but also society's welfare, economic and social progress, and quality of life. Moreover, big data can be used to improve sustainability in many aspects. Specifically, the importance of analyzing both terms together is relevant because it can help public administrators and practitioners to develop Big Data applications, with sustainability concerns, and to better understand the conditions under which the implementation of Big Data can create benefits to improve the environmental, societal, and economic sustainability. For example, Big Data can benefit public organizations and address societal issues, as well as can help firms outperform their competitors [50]. In addition, the introduction of sustainable practices and strategies can also reduce costs, improve the organizations' environmental, social and governance ratings or even lead them to build better performance and competitive advantage (as it can lead to more operational efficiency, or to offer more value to customers, resulting in higher number of customers, better customer engagement, or even more turnover). Moreover, the sustainable practices can help managers and practitioners develop and make better decisions; their analysis and study can help researchers to better understand the causal mechanisms that lead companies to embrace sustainability in a successful way [51].

Consequently, given the lack of bibliometric analyses on BD&S, and given the need for this study, our research attempts to fill this gap by carrying out a bibliometric analysis that addresses this issue.

The paper is structured as follows: the next section discusses the concepts of Big Data and Sustainability as well as some of their interrelationships; Section 3 describes the data sources and bibliometric methodologies used to delve into the interrelationships; Section 4 presents the results of the analysis, visualizes the data, and presents the main findings; Section 5 provides discussion of the results, and, finally, Section 6 offers the main research conclusions.

## 2. Literature Review

The evolution from Web 2.0 through the Semantic web to the Ubiquitous Web of today, which involves the interaction between technology and Big Data, must consider the key role of sustainability for the development of any field. First, it is necessary to understand the new transformations and why Big Data is so important. The new organizational context has been increasingly defined by the advancement of Big Data, which has been triggered by the development of information technologies, social media, the Internet, Internet of Things, and cloud computing. As a result of these digital and technological innovations, an explosive growth of data in almost all industries and business areas is occurring [52]. In addition, Big Data has been increasingly recognized as one of the critical resources for the development and success of organizations. These developments are reinforced by big data-related tools and activities, such as learning machines and artificial Intelligence, robots, smart communication devices, business intelligence and analytics. Overall, those advancements represent the so-called Ubiquitous Era, which is conceptualized in this paper as a new epoch in which companies have to provide customized solutions to address the needs of customers in their specific context. More specifically, the concept is related to a global computing environment, which can provide users with "seamless and invisible access to computing resources" [53,54] (p. 421).

Big Data has proved essential not only for an improved customer service, but also for an enhanced internal management of all types of organizations and geographic areas. As such, big data has also incited academic interest, judging by the growth pattern of publications on Big Data since the beginning of the last decade [26] (p. 1179). The topic has rapidly evolved into a hot research area, attracting widespread attention from academia, industry, and governments around the world [52] (p. 59) across a wide range of disciplines [15]. The academic research on big data is particularly important in the business field [55], as Big Data is used in all areas of industry and business functions nowadays. The term refers to a large mass of digital data collected from various sources [15]. This rich data shows more

details about the behaviors, activities, and events happening, whereby big data analytics gives access to the “variety and different types of data, from huge resources, with less response time” [56] (p. 111). It integrates the physical world (obtaining data through sensors, scientific experiments, and observations), human society (gathering data through social networks, the internet, and other mechanisms), and Cyberspace [52].

But what is the relationship between these processes and sustainability? Sustainability has developed as an important, growing research discipline and is a policy-relevant issue for businesses and public organizations. Based on the need to preserve limited resources, and, in line with the Brundtland report and Garrigos et al. [2], sustainability is defined as those actions and developments that meet the needs of today’s societies, organizations, and people, without having a negative impact on the environment, ecology, society, landscape, culture, or heritage, and without compromising the prosperity and well-being of future generations. The field incorporates diverse perspectives, including natural, cultural, educational, human, political, institutional, and socio-economic ones and, essentially, represents a balance between economic, ecological, and social dimensions. In this regard, recent literature reveals economic (performance), environmental (energy, products, and services), and social (employment, local community) sustainability categories associated with corporate sustainability [57]. In addition, bringing economic growth and wealth should go hand in hand with environmental and societal concerns that affect key stakeholders, both internal and external. These stakeholders include people, governments, organizations, and industries, and their various purposes related to increasing wealth, quality of life, economy, and competitiveness. In fact, as Sanchez Planelles et al. [51] state, sustainability practices, from a managerial point of view, have to respond to new challenges and stakeholder pressures. In addition, these authors observe different concepts of sustainability from a corporate sustainability perspective, including holistic sustainability, sustainable business models, sustainable methodologies, sustainable operations, and sustainability-oriented innovation.

In this context, Big Data management, combined with technological development, is essential to reach the sustainability goals in the organizational world as digitalization widely impacts business and society [58]. More specifically, it should be noted that the transformations inherent to Big Data, and the development of the Ubiquitous Web, are changing the roles of both customers and organizations through sophisticated, open innovation techniques, co-design, co-creation, co-working, co-marketing, and crowdsourcing (“the act of taking a specific task and outsourcing it to a large group of people over the internet, through an open call” [59] (p. 95). As a result, there is increased customer involvement, as consumers become protagonists, in the design and creation of customized and personalized products and services [60–63].

However, its application is more extensive. Understanding big data and its consequences is vital to: create and develop new business models; facilitate management (e.g., business intelligence) and make better business decisions; design, develop, deliver, and sell better, and more personalized, services and products (as it offers information about customers’ preferences, feedback about organizations’ products and services performance, and insight about emerging trends [50]; create, optimize and improve business processes in terms of management and distribution; improve overall and employee productivity, efficiency and reduce costs; increase connectivity and self-learning; improve and streamline internal and external operations and communications (e.g., by incorporating knowledge of user interaction in social networks; improvements in interactions with users (B2C) and businesses (B2B)); improve public and private cooperation, transparency and participation; assist customers in real life (e.g., to interact with customers at any point in their life or experience, to provide a pocket assistant, or to understand specific customer needs through the use of artificial intelligence and smart devices); plan, predict and monitor various developments and effects surrounding organizations and geographical areas; develop green innovation and the circular economy; or to improve marketing (helping to understand, satisfy and improve customer loyalty and engagement). All this leads to enhancing the competitiveness and sustainability of organizations. For instance, green management prac-

tices related to the integration of big data technology, can improve banks' environmental and financial performance [64]. Big data is also able to generate valuable information and insights, helping to capture value about customers, and impacting the long-term firm performance [50]. Moreover, Big Data can improve societal and environmental sustainability in supply chains [65] by reducing, for instance, slavery or pressures on the consumption of natural resources through the improvement of visibility and coordination among supply chain partners.

Furthermore, due to COVID-19 the use of Big Data and new technologies is essential not only to control and reduce the impact of the pandemic but also to improve people's quality of life, as well as the business ecosystem and the environment. In this regard, the use of Big Data is essential to predict, plan, manage, and control environmental processes, which derives from a discipline called Environmental Informatics [66]. Combining Artificial Intelligence (AI), Geographic Information Systems (GIS), modelling, simulation, and user interfaces is a key tool for environmental sustainability. These can also be used to improve green infrastructure and practices to optimize energy, water, waste, pollution, and other externalities, and to improve resource usage and thus reduce negative environmental impacts. For instance, big data provide new and powerful ways of studying and improving coupled environmental, social, and economic systems to enhance urban sustainability [47]. Big data can be used to strengthen the sustainability of healthcare project financing by improving the quality and timeliness of information [67]. The capability to analyze and visualize large datasets in a rapid timeframe, which is possible thanks to Big Data, can help to improve the conservation and sustainability of the environment, as well as mitigate environmental declines, by allowing scientists to discover, analyze, and better understand environmental changes at micro to global scale [68].

However, the development of Big Data also presents some negative consequences for environmental sustainability, as it can lead to overexploitation of resources needed to develop and implement new technological innovations. An increased demand for scarce resources, in turn, requires more energy production, which results in a higher pollution rates and other environmental impacts. The sustainability of organizations or the functioning of the economy can also be negatively affected, as it can lead to monopolistic practices. More specifically, the leading technological companies can dominate a wide range of sectors as they control big data, while small and medium enterprises do not have enough knowledge, access to large volume of data, technological resources and financial capability to implement innovations. It can also transform social and cultural structures, affect employment, people's privacy, safety and security, the income distribution, and even affect the concept of society and its governance as we know it today.

### 3. Materials and Methods

This research uses the WoS Core Collection database as a data source. The choice of this database is justified by the fact that it is one of the two databases with the highest academic recognition, together with Scopus. In this study, Scopus was discarded as WOS is more restrictive in terms of the quality of the journals it includes [69]. The option of Google Scholar was not deemed suitable due to existing unreliability concerns [70]. Thus, our approach follows previous studies that have used WoS [2,71].

In order to evaluate only papers that jointly analyze the topics of "Big data" and "sustainability," we searched for all the papers that use both keywords simultaneously in the entire database (including all its sub databases, knowledge areas and categories [72]. However, the population only comprises documents catalogued as articles, reviews, letters, and notes [2]. Documents from 2011 (the first year in which a document on both topics appears), until 31 December 2020 are included. Data collection took place in February 2021. The final sample consists of 726 documents.

The study uses the VOSviewer software [73] to map and graphically illustrate the data. This tool is widely used in the literature to visualize the structure and networks of authors, journals, universities and countries [72]. In addition, the study makes use of the most

broadly used analyses in the bibliometric literature [2]: keywords co-occurrence, co-citation (two papers cited by the same document) [74]; bibliographic coupling (one document is cited by two documents) [75] and co-authorship [71]. The research performed hierarchical cluster analyses.

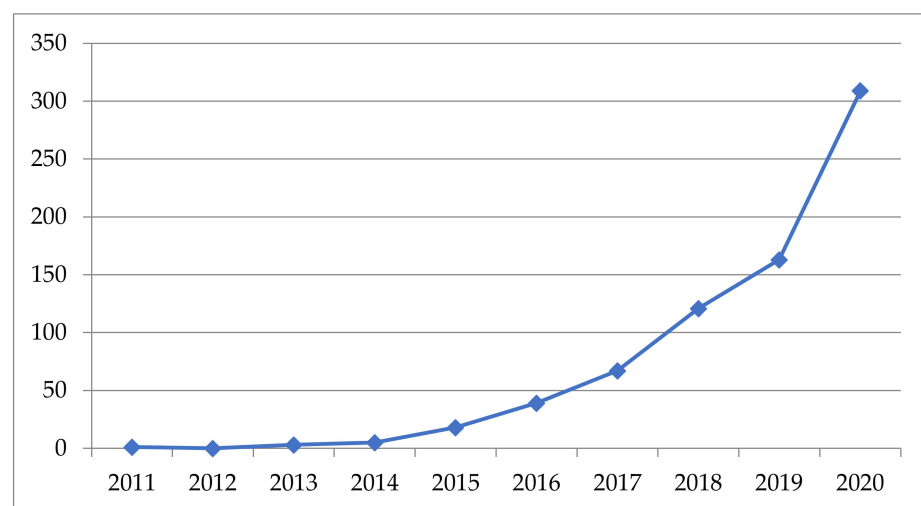
Moreover, the most popular bibliometric indicators were used including the total number of published articles (to assess productivity); the total number of citations (to observe the relevance of an author, an institution or a country [69,71]); the h-index (indicating the authorship of h documents cited at least h times [15] to analyze the quality of a group of documents [76]; the number of documents above a number of citations threshold (to observe influence) [2]; the impact factor reflected in the WoS (to show the dissemination power of a source) [71]; the citation/document ratio (to measure the impact of each article).section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

#### 4. Results

In this section the results derived from seven analyses are presented. Firstly, the state of the art, the research progress on BD&S and the citation structure of the articles is described. Next, the most cited papers on BD&S are analyzed. The third subsection contains the results related to the top journals. The fourth part focuses on the co-occurrence study of keywords associated with BD&S. This is followed by the analysis of co-citation of references, journals, and authors in the BD&S literature. The sixth subsection focuses on the bibliographic coupling of authors. Finally, the study reveals the co-authorship clusters of countries and institutions.

##### 4.1. Current Situation and Evolution of the Literature on Big Data and Sustainability

The first paper on Big Data and Sustainability, published in the Web of Science (WoS) core collection, analyses spatiotemporal patterns in criminal offense records [77]. Although there were only 3 published articles in 2013 and 5 in 2014, the growth of the academic literature on the topic has been exponential since then. Thus, in 2018, 121 articles were published, reaching 309 in 2020, with an increase of almost 90% in the last year. Figure 1 indicates the annual trend in terms of publications.



**Figure 1.** Annual publications in WoS on Big Data and Sustainability. Data from the WoS.

The importance of the articles on Big Data and sustainability is highlighted by the number of citations they have received. The most cited paper in this field is Tao et al. [78], which received 346 citations in WoS in just over two years. The articles by Bibri and

Krogstie [79] and Nuaimi et al. [80] have also received more than two hundred citations from papers indexed in WoS.

Table 1 shows the general citation structure of BD&S publications. The data indicates that only the first two articles obtain more than 250 citations (0.28%), although 29.48% of them exceed 10 citations, despite the novelty of the topic. In addition, the h-index [76], which is indicative of the holistic observation of the field [72], for all articles related to BD&S, is 49 (49 articles have 49 or more citations). In addition, the whole sample of articles is cited more than 9500 times in WoS, with an average of 13.25 citations per item.

**Table 1.** General citation structure of BD&S publications.

N° of Citations	N° of Articles	Accumulated N° of Articles	% Articles	% Accumulated Articles
≥250	2	2	0.28	0.28
≥200	1	3	0.14	0.41
≥100	12	15	1.65	2.07
≥50	31	46	4.27	6.34
≥25	61	107	8.40	14.74
≥10	107	214	14.74	29.48
<10	512	726	70.52	100.00
Total	726			

Source: Own elaboration based on WoS 2021.

#### 4.2. Top Cited Articles on BD&S

Table 2 contains the most important documents in the BD&S research area. Specifically, the data shows the characteristics of the top 20 articles that have receive the greatest number of citations. This analysis is very relevant, since this data reveals not only the quality of the documents, but also their popularity and influence in the field [72].

**Table 2.** Top 20 papers on BD&S with the most citations.

R	Journal	TC	Article	Authors	Year	CY
1	IJMT	346	Digital twin-driven product design, manufacturing and service with big data	Tao, Fei; Cheng, Jiangfeng; Qi, Qinglin; et ál.	2018	114.33
2	SCS	266	Smart sustainable cities of the future: An extensive interdisciplinary literature review	Bibri, Simon Elias; Krogstie, John	2017	65.75
3	JISA	215	Applications of big data to smart cities	Al Nuaimi, Eiman; Al Neyadi, Hind; Mohamed, Nader; et ál.	2015	35.83
4	AC	174	Enhancing environmental sustainability over building life cycles through green BIM: A review	Wong, Johnny; Kwok Wai; Zhou, Jason	2015	28.50
5	JCP	155	The role of Big Data in explaining disaster resilience in supply chains for sustainability	Papadopoulos, Thanos; Gunasekaran, Angappa; Dubey, Rameshwar; et ál.	2017	38.75
6	PSEP	142	Sustainable Industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives	Kamble, Sachin S.; Gunasekaran, Angappa; Gawankar, Shradha A.	2018	46.33
7	ISJ	129	Big Data Meet Green Challenges: Big Data Toward Green Applications	Wu, Jinsong; Guo, Song; Li, Jie; et ál.	2016	25.80
8	S	129	What Drives the Implementation of Industry 4.0? The Role of Opportunities and Challenges in the Context of Sustainability	Mueller, Julian Marius; Kiel, Daniel; Voigt, Kai-Ingo	2018	42.33
9	BDR	122	Big Data Analytics for Dynamic Energy Management in Smart Grids	Diamantoulakis, Panagiotis D.; Kapinas, Vasileios M.; Karagiannidis, George K.	2015	20.33
10	AAAJ	119	Achieving the United Nations Sustainable Development Goals: An enabling role for accounting research	Bebbington, Jan; Unerman, Jeffrey	2018	39.67

Table 2. Cont.

R	Journal	TC	Article	Authors	Year	CY
11	JDMM	119	Big data analytics for knowledge generation in tourism destinations—A case from Sweden	Fuchs, Matthias; Hoepken, Wolfram; Lexhagen, Maria	2014	17.00
12	AFM	117	Wearable and Miniaturized Sensor Technologies for Personalized and Preventive Medicine	Tricoli, Antonio; Nasiri, Noushin; De, Sayan	2017	29.25
13	ITCC	113	A Secure Cloud Computing Based Framework for Big Data Information Management of Smart Grid	Baek, Joonsang; Quang, Hieu Vu; Liu, Joseph K.; et al.	2015	18.83
14	ITSG	111	Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges	Wang, Yi; Chen, Qixin; Hong, Tao; et al.	2019	55.50
15	SCS	111	The IoT for smart sustainable cities of the future: An analytical framework for sensor-based big data applications for environmental sustainability	Bibri, Simon Elias	2018	37.00
16	IJIM	95	Big data reduction framework for value creation in sustainable enterprises	Rehman, Muhammad Habib Ur; Chang, Victor; Batool, Aisha; et al.	2016	19.00
17	CIE	94	Big data analytics in supply chain management between 2010 and 2016: Insights to industries	Tiwari, Sunil; Wee, H. M.; Daryanto, Yosef	2018	31.33
18	JCP	93	Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties	Wu, Kuo-Jui; Liao, Ching-Jong; Tseng, Ming-Lang; et al.	2017	23.25
19	ISJ	91	Big Data Meet Green Challenges: Greening Big Data	Wu, Jinsong; Guo, Song; Li, Jie; et al.	2016	18.20
20	SCS	90	Can cities become smart without being sustainable? A systematic review of the literature	Yigitcanlar, Tan; Kamruzzaman, Md.; Foth, Marcus; et al.	2019	45.00

Source: Own elaboration based on WoS 2021. R: Ranking; TC: Total Citations; CY: Citations per year. SC: Science; IJTM: International Journal of Advanced Manufacturing Technology; SCS: Sustainable Cities and Society; JISA: Journal of Internet Services and Applications; AC: Automation in Construction; JCP: Journal of Cleaner Production; PSEP: Process Safety and Environmental Protection; ISJ: IEEE Systems Journal; S: Sustainability; BDR: Big Data Research; AAJ: Accounting Auditing & Accountability Journal; JDMM: Journal of Destination Marketing & Management; AFM: Advanced Functional Materials; ITCC: IEEE Transactions on Cloud Computing; ITSG: IEEE Transactions on Smart Grid; SCS: Sustainable Cities and Society; IJIM: International Journal of Information Management; CIE: Computers & Industrial Engineering.

The article by Tao et al. [78] leads not only in terms of citations in WoS (346), but also in regard to the number of citations per year (115.33). This article studies “how to generate and use converged cyber-physical data to better serve product lifecycle, so as to drive product design, manufacturing, and service to be more efficient, smart, and sustainable.” The second most cited paper is Bibri and Krogstie [79] with 266 citations, which is also the second most cited paper per year (66.5) in the ranking. This paper provides an overview of the field and existing work on smart (and) sustainable cities and reveals numerous research opportunities in the area of smart sustainable cities. The article by Al Nuaimi et al. [80] is third in the ranking of citations (215), although it occupies a lower position when considering the number of citations per year (35.83). This paper, also focused on the study of smart cities, analyses the use and implementation of big data applications to support smart cities in order to improve sustainability and the living standards (e.g., to improve the performance of health, transportation, energy, education, and water services). Finally, the paper by Wang et al. [81] should be pointed out, which, although occupying a discrete position in the ranking of total amount of citations (111 citations), is ranked third as for citations per year with 55.5 annual cites in WOS. This paper reviews smart meter data analytics, highlighting the relevance of the massive use of this data to promote and enhance the efficiency and sustainability of the power grid.



### 4.3. Leading Journals in BD&S

The three main publication categories within the WoS, are Green Sustainable Science Technology (211 papers, 29.06%), Environmental Sciences (193 papers, 26.58%), and Environmental Studies (145, 19.45%). Analyzing the sources of the 726 papers in our sample, these were published in 370 journals, only 47 of which have published more than 3 papers, 19 published more than 5, and only 6 more than 10 documents on BD&S.

Focusing on these sources, 234 documents (22.23% of the BD&S sample) appeared in the top 10 journals (Table 3). However, only the first two journals account for more than 20% of the articles published on the topic: these are Sustainability accounting for 15.29% of the total publications, Journal of Cleaner Production representing 5.51%, and in third place Sustainable Cities and Society with 2.48%. The H-index is also led by these three sources, although in this case Journal of Cleaner Production (17) ranks first, followed by Sustainability (13) and Sustainable Cities and Society (7).

**Table 3.** Top 19 journals per number of articles on BD&S.

R	Journal	APBS	H-BS	TAP	TCBS	ACBS	PCBS	%APBS	IF	≥200	≥100	≥50	≥20
1	S	111	13	28,156	840	680	7.57	0.39	2.58		1	5	9
2	JCP	40	17	20,733	885	778	22.13	0.19	7.25		1	3	15
3	SCS	18	7	2476	600	462	33.33	0.73	5.27	1	2	4	5
4	IA	13	5	43,099	135	135	10.38	0.03	3.75				2
5	ASTI	11	1	87	2	2	0.18	12.64	-				-
6	PPC	11	6	975	87	78	7.91	1.13	3.61				1
7	E	8	6	20,164	162	158	10.88	0.82	2.70			2	5
8	JEIM	8	2	360	8	7	1.00	2.22	2.66				0
9	CIE	7	5	3849	279	263	39.86	0.18	4.14			3	4
10	IJPE	7	5	3192	157	153	22.43	0.22	5.13			1	3
11	MD	7	3	1326	40	39	5.71	0.53	2.72				1
12	FGCS	6	3	3589	106	106	17.67	0.17	6.13			1	2
13	ISF	6	2	807	31	30	5.17	0.74	3.63				-
14	CI	5	4	959	128	125	25.60	0.52	3.95			1	2
15	IJIM	5	5	1065	248	234	49.60	0.47	8.21			2	4
16	IJLM	5	3	398	65	60	13.00	1.26	3.33				2
17	IJPR	5	4	4655	111	103	22.20	0.11	4.58				3
18	SS	5	2	656	17	17	3.40	0.76	5.30				-
19	TFSC	5	4	2774	192	166	38.40	0.18	5.85			2	4

Source: Own elaboration based on WoS 2021. R: Ranking; H-BS: indicates the H index in the area of Big Data & Sustainability; APBS: Articles published on BD&S; TAP: Total Articles published; TCBS: Total citations in BD&S; ACBS: Articles in which is cited in BD&S; PCBS: Average of cites by articles in BD&S.; %APBS: Percentage of articles published in BD&S (APBS/TAP); IF: Impact Factor; ≥200. ≥100. ≥50 and ≥20: articles with more of 200.100.50 and 20 citations; S: Sustainability; JCP: Journal of Cleaner Production; SCS: Sustainable Cities and Society; IA: IEEE Access; ASTI: Advances in Science Technology Innovation; PPC: Production Planning Control; E: Energies; JEIM: Journal of Enterprise Information Management; CIE: Computers Industrial Engineering; IJPE: International Journal of Production Economics; MD: Management Decision; FGCS: Future Generation Computer Systems The International Journal of Science; ISF: Information Systems Frontiers; CI: Computers in Industry; IJIM: International Journal of Information Management; IJLM: International Journal of Logistics Management; IJPR: International Journal of Production Research; SS: Sustainability Science; TFSC: Technological Forecasting and Social Change.

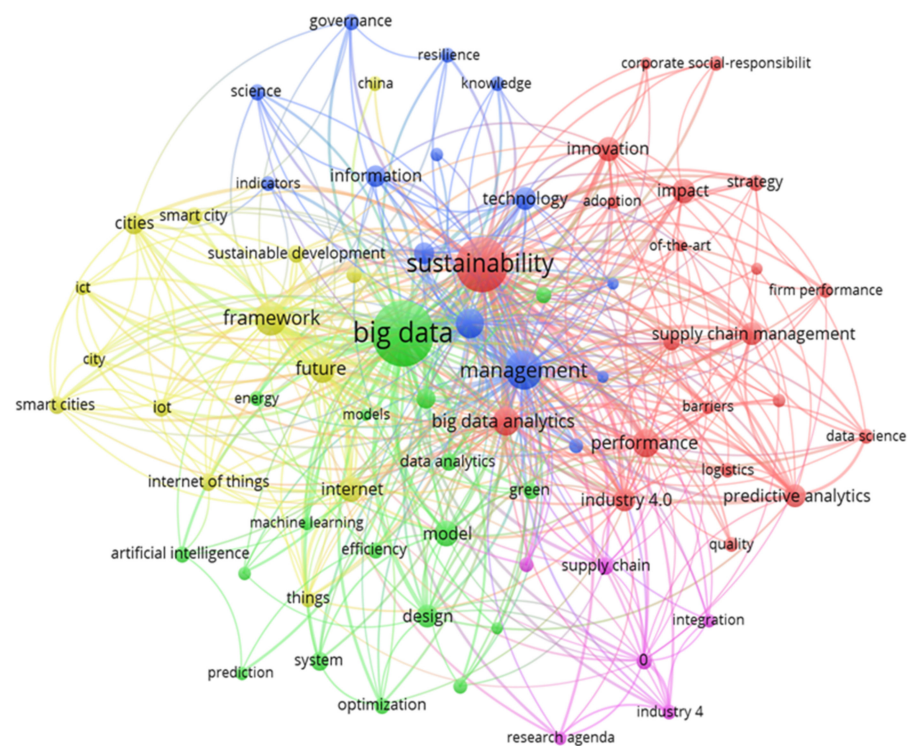
Analyzing further the journals in Table 3 and focusing on those that have dedicated the largest amount of their published articles to the BD&S topic since 2011 (the first year in which a publication on the topic appeared), Advances in Science Technology Innovation comes first with 12.64% of its publications dedicated to the topic. The Journal of Enterprise Information Management, and International Journal of Logistics Management are the next in the ranking, although these sources have only devoted 2.22% and 1.25% respectively of their publications to BD&S.

As for the sources with the highest number of citations per published article on BD&S, the ranking is led by the International Journal of Information Management (49.6 citations on average), followed by Computers Industrial Engineering (39.86), Technological Forecasting and Social Change (38.40), and Sustainable Cities and Society (33.33) in fourth position.

#### 4.4. Keywords Analysis

In this section, keywords are analyzed in order to get a glimpse of their importance and distribution, which, in turn, represent the state of the art and the main trends in the BD&S area. More specifically, co-occurrence analysis (which studies the number of articles in which two keywords appear together) of all keywords was performed. The analysis of the 726 documents in sample, through the VOS viewer program, indicates the existence of 3802 keywords.

Figure 2 shows the most relevant keywords and the size of the nodes (the larger the keyword and the node, the greater the number of articles that contain the keyword), as well as the relationships between them. The thickness of the relationship line indicates the frequency of the joint co-occurrence of two keywords, while the relative strength of the relationship between two keywords, compared to others, is displayed by the smaller distance between them. The nodes are colored to indicate the different clusters and groups of keywords. Using a threshold of fifteen co-occurrences, Figure 2 shows the 73 most frequent keywords grouped in 5 clusters. The three main keywords are “big data”, “sustainability” and “management.”



**Figure 2.** Co-occurrence network of all keywords of BD&S-related publications. A threshold of 15 occurrences is used, which yielded 73 keywords.

However, the most relevant cluster is the one colored in red and labelled “sustainability” with 21 items. The network includes, among others, the keywords “big data analytics”, “performance”, and “innovation”, listed as the sixth, seventh, and twelfth most frequent keywords. The second cluster, in green, contains 18 items, dominated by the keyword “big data”, and also including “model” and “design” as relevant keywords. The third cluster, displayed in blue, is labelled as “management” and contains 14 items among which “challenges”, “technology”, and “analytics” should be highlighted. The fourth cluster, in yellow, also comprises 14 items, being “framework”, “future”, and “internet” the fourth, the eighth and the tenth most frequently used keywords. Finally, the fifth cluster colored in purple contains six items, only two out of which are among the 30 most frequent keywords: “industry 4.0” and “supply chain.” Table 4 shows the top 30 keywords, their frequency, and total link strengths.

**Table 4.** All keywords' co-occurrences of BD&S-related publications.

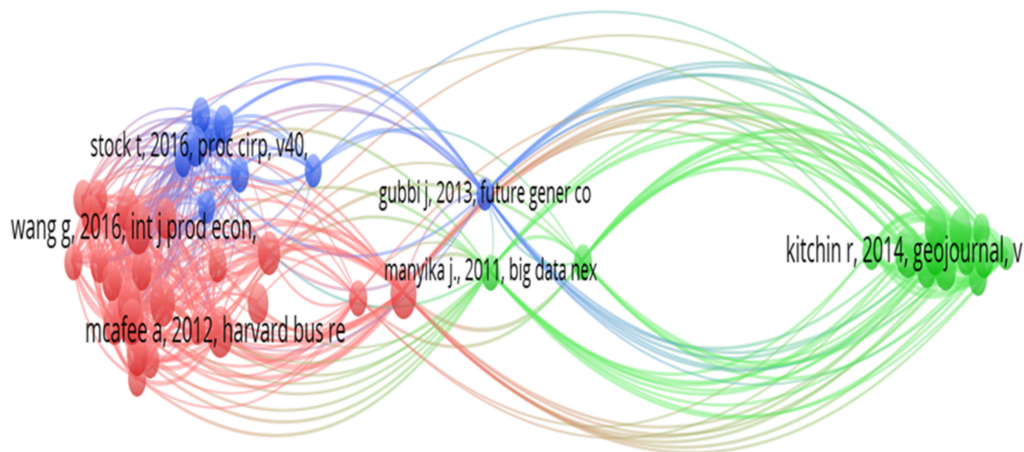
R	Keyword	Oc	Co
1	Big Data	407	1628
2	Sustainability	289	1338
3	Management	143	733
4	Framework	110	588
5	Challenges	80	458
6	Big Data Analytics	77	411
7	Performance	73	397
8	Future	70	424
9	Model	62	266
10	Internet	61	362
11	Impact	56	272
12	Innovation	56	303
13	Design	52	265
14	Supply Chain Management	49	292
15	Technology	49	251
16	Industry 4.0	48	318
17	Predictive Analytics	48	301
18	Analytics	45	200
19	Cities	44	206
20	Information	43	197
21	Systems	40	211
22	System	35	151
23	Smart Cities	32	145
24	Internet of Things	31	167
25	Sustainable Development	31	167
26	Data Analytics	30	169
27	Supply Chain	30	174
28	Circular Economy	29	195
29	IoT	29	183
30	Smart City	9	23

Note: R: Rank; Oc: All keyword occurrences; Co: All keyword co-occurrences link.

#### 4.5. Reference, Journal and Author Co-citation Analysis

The study of co-citations indicates when two elements (author, journal, or document) are cited simultaneously by a third document, thus appearing together in a new bibliographic list [74]. This subsection analyzes three elements: references, journals, and authors. In doing so, the characteristics, development, structure, and relationships of BD&S research is revealed, through the analysis of relationships within the current sample.

Figure 3 illustrates the main co-cited papers. A more detailed analysis shows that in terms of the number of citations the study of Wang et al. [82] and McAfee et al. [83] share the first position in the ranking with 47 citations each by the 726 papers in the BD&S sample. In this ranking the paper by Kitchin [84] ranks third with 45 citations. As for the ranking based on the link strength, the first and third positions are still occupied by Wang et al. [82] and Kitchin [84], although the second one belongs to Gunasekaran et al. [85] with respective values of 263, 237, and 232. It should be noted that these papers are not necessarily included in the sample of 726 papers on BD&S, but that they are co-cited by them.



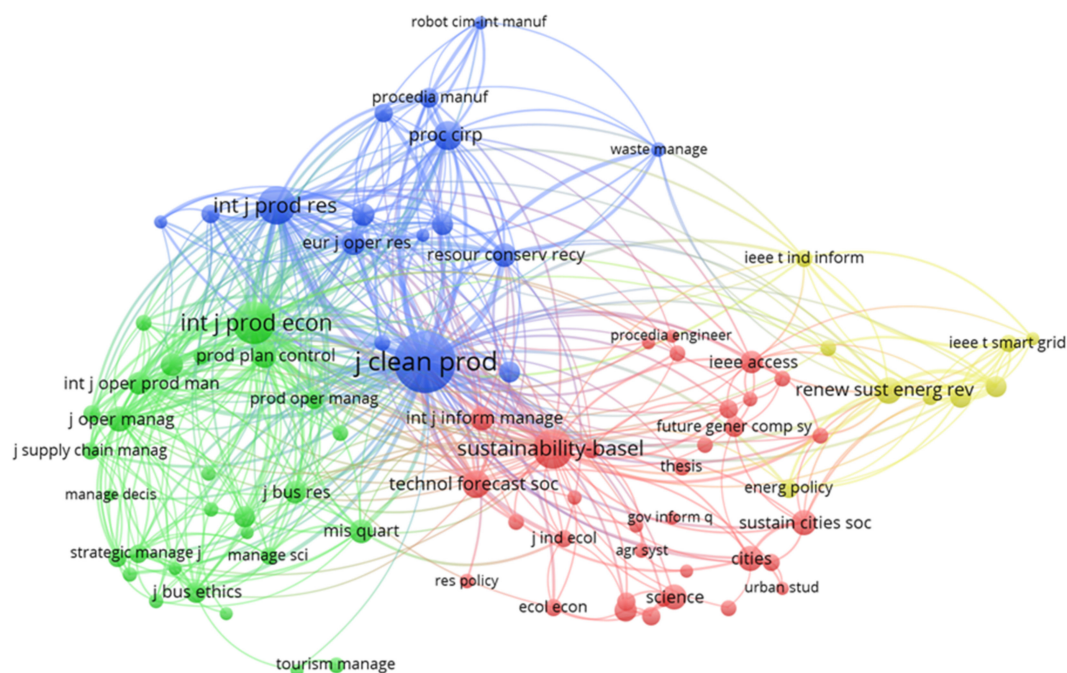
**Figure 3.** Co-citation of cited references on BD&S: 47 references of the 44,856 cited references that meet the citation threshold of a minimum number of citations of a cited reference of 20.

The documents by Wang et al. [82], McAfee et al. [83], and Gunasekaran et al. [85] represent the largest cluster, which is colored in red, in Figure 3 and encompasses 24 items. This network has a predominantly management focus related to Big Data. More specifically, these three papers mainly focus on Big Data management for value chain logistics management, as well as supply chain networks for sustainability. The rest of the papers in this cluster have similar characteristics. The second cluster, displayed in green, contains 14 items and is led by the study of Kitchin [84]. This cluster has a more targeted approach to sustainability issues, considering the geographic scopes and smart areas (including cities). For instance, Kitchin [84] focuses on the use of Big Data for smart urbanism. A large part of the research gathered in this cluster addresses the sustainable management of cities, and the so-called smart sustainable city development. The emphasis is on the management of areas and destinations, rather than the management of specific organizations and their value chains. In addition, while the previous network is mainly centered on the technological side of big data, this cluster emphasizes the sustainability and the environmental aspects. Finally, the blue cluster, encompassing only nine items, is closely related to the first one. This last cluster is dominated by Stock and Seliger [86], although the article did not perform well in the citation (30 citations) nor in the link strength ranking (27th position). Stock and Seliger [86] focus on sustainable manufacturing in the industry 4.0. In contrast to the first cluster, which also addresses technology management, the focus here is on the technological rather than the management aspects (e.g., Internet of things, industry 4.0, etc.)

Once the reference co-citation results are presented, the journal co-citation network on BD&S is analyzed (Figure 4). As in the previous analysis, the nodes in the illustration show the importance and number of cited items, while the distances between them indicate the frequency of co-citations. The analysis uncovers the existence of four networks of co-cited journals. The first network, displayed in red, contains 34 journals, with *Sustainability* (675 citations, 22,698 link strength) and *Technological Forecasting & Social Change* (340 citations, 13,210 link strength) the most relevant ones. These journals are ranked fourth and ninth in terms of link strength, but third and sixth as per number of citations. This cluster is focused on sustainability, looking specifically at the sustainability of cities (as it includes the journals *Cities and Sustainable Cities and Society*) and the field of empirical science, including *Sciences and Nature*.

The second cluster, in green, contains 29 items, among which the *International Journal of Production Economics* (762 citations, 40,286 link strength), *Journal of Operations Management* (270 citations, 16,311 link strength), and *Production Planning & Control* (239 citations, 13,938 link strength) are the most relevant. These journals are ranked 2nd, 6th, and 7th in terms of link strength, but in terms of citations they occupy the 2nd, 7th, and

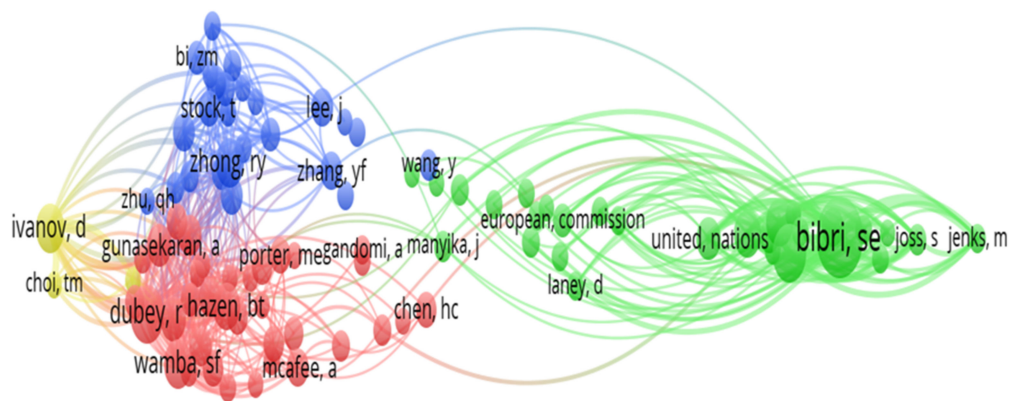
12th position respectively. This cluster is a predominantly management-oriented, as it also includes journals such as International Journal of Operations & Production Management (223 citations, 13,014 link strength), Supply Change Management (216 citations, 11,367 link strength), and MIS Quarterly (241 citations, 9170 link strength). However, the prevailing management area is information systems management, associated with decision making. This cluster shares information management content with the previous cluster, and also addresses operations management issues as the next one.



**Figure 4.** Journal co-citation network on BD&S: 88 main journals of the 17,913 cited sources by the 726 documents regarding BD&S that meet the threshold of a minimum number of citations of a cited source of 70.

The third largest cluster, the blue one, includes 16 items, the Journal of Clean Production (1659 citations, 67,793 link strength), International Journal of Production Research (650 citations, 38,520 link strength), and Procedia CIRP (359 citations, 18,431 link strength) being the most important ones. While they rank 1st, 3rd, and 5th in terms of link strength, they hold the 1st, 4th, and 5th position in the ranking per number of citations. This cluster also focuses on the management area, but in this case, it is centered on production management, although it shares this topic with the previous one. Among the most relevant journals pertaining to this cluster are Resources Conservation & Recycling (253 citations, 13,410 link strength), Computers & Industrial Engineering (233 citations, 13,087 link strength), and European Journal of Operational Research (243 citations, 12,523 link strength). Finally, the last cluster encompasses only 9 items, among which is Renewable & Sustainable Energy Reviews (293 citations, 11,443 link strength). This network is characterized by its focus on the application of new technologies in the energy field.

The last co-citation analysis aims to reveal the main authors on BD&S. Figure 5 unveils a structure of four networks of authors and the relationships among them. The cluster with the greatest number of authors contains 35 items and is depicted in red. The most cited scholar is Dubey (120 citations), followed by Wamba (87 citations) and Hazen (76 citations), who occupy the 2nd, 5th, and 8th positions in the citation ranking, respectively. This network includes another six authors in the top 20, namely Song (70 citations), Gunasekaran (65 citations), Seuring (55 citations), Hair (54), Wang (53 citations), and McAfee (52 citations), in the 10th, 11th, 13th, 14th, 16th, and 17th positions. The studies conducted by these researchers focus on big data in general, as well as its application in logistics and supply chain management, as already observed in the co-citations of references analysis.



**Figure 5.** Author co-citation network on BD&S: 92 authors, of the 32,765 cited authors, which meet the threshold of a minimum number of citations of a cited author of 25.

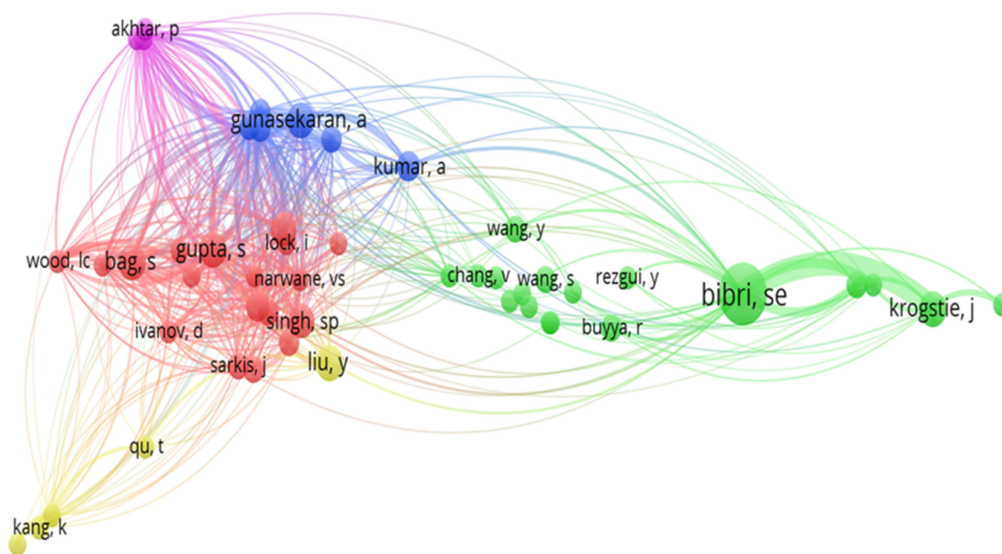
The second most relevant cluster, displayed in green, contains 29 authors, among which Bibri (226 citations) and Kitchin (117 citations) occupy the 1st and the 3rd positions in the ranking. Other prominent scholars in this cluster are Batti (78 citations), Yigitkanlar (75 citations), United Nations (64 citations) and Angelidou (53 citations), in the 7th, 9th, 12th, and 15th positions, respectively. The research developed by Bibri and Kitchin is related to smart sustainable cities and smart urbanism. This highlights, once again, the importance of this topic, which was also identified in the analysis of co-citations of references.

As for the third cluster, encompassing 25 items, the most distinguished scholar is Zhong (81 citations), who is the 6th most cited author. Although this cluster does not contain other authors in the top 10, it includes Govindan (51 citations), Zhang (51 citations), and Jabbour (49 citations). These authors have researched the field of sustainable manufacturing in industry 4.0.

Finally, the last cluster, in yellow, includes three authors, among which Ivanov (87 citations) stands out as the 4th most cited author. It should be noted that this cluster was not revealed in the analysis of co-citations of references. Ivanov's research addresses the combination of digital technology, smart factory industry 4.0, and value chain management from a mathematical and engineering perspective. The position of the cluster in the figure, far from the second one and in the middle, between the first and the third one, is due to its technological approach.

#### 4.6. Bibliographic Coupling of Authors

Bibliographic coupling complements the previous analysis of the relationship among authors by counting the number of references that a group of documents have in common (documents A and B are coupled if they both cite another article C) [75]. Figure 6 shows the results of this analysis, which indicates the existence of five main groups of authors. Looking at the link strength ranking, Bibri (31 papers and 3941 link strength), Krogstie (7 papers and 2991 link strength), and Gunasekaran (7 papers and 2043 link strength) occupy the first three positions. These authors are also leading the ranking in terms of total citations, although the order is as follows: Gunasekaran (505 citations), Bibri (500 citations), and Krogstie (318 citations). Drilling down into the structure of clusters, the first of them, in red, and with 18 members, is anchored by Bag and Gupta, ranked 4th and 8th according to the link strength criterion, with 7 documents each. Jabbour is another prominent scholar, who appears with two entries, and such is Hazen, who is ranked 8th in the citation ranking with six papers. The research of these authors is focused on industry 4.0.



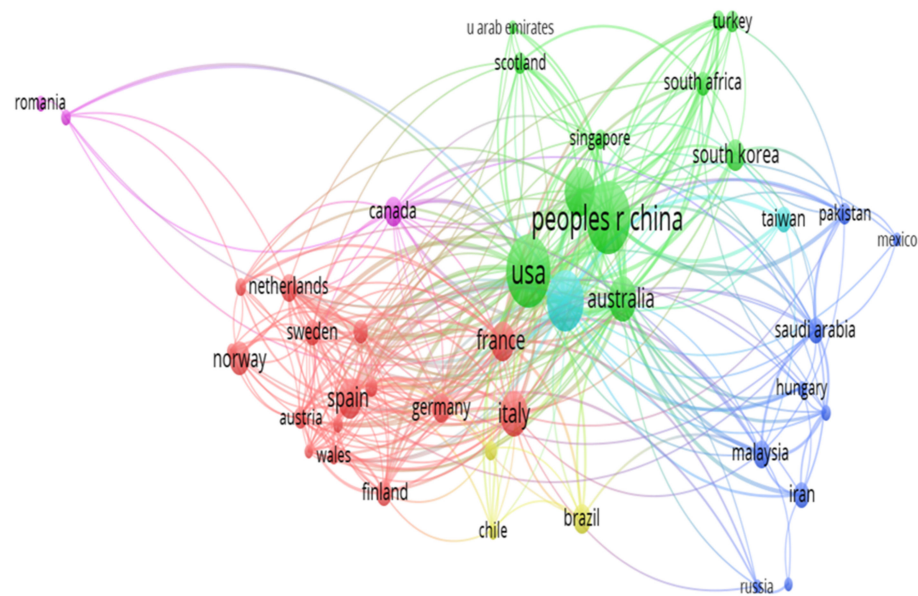
**Figure 6.** Bibliographic coupling of authors: 55 authors of the 2343 authors which meet the threshold of a minimum number of documents per author of 3.

The second cluster, depicted in green, contains 18 authors, the leading among them being Bibri and Krogstie. A close relationship between both scholars is observed, which is explained by their frequent co-authorship (e.g., [79]). Another distinguished researcher in this cluster is Wu, who ranks 6th per number of citations. As mentioned earlier, the research area of these authors is smart sustainable cities. The third network, in blue, contains six researchers, Gunasekaran being the most prominent among them. Childe, Dubey, and Papadopoulos, with four articles each, occupy the 5th, 6th, and 7th positions, respectively, in the link strength ranking. This group of researchers, which also includes Kamble, investigate Big Data analytics in logistics and supply chain management. Next is the yellow cluster, comprising five authors and anchored by Y. Liu, who, although in a low position, according to the link strength criterion, receives 58 citations from his seven papers. His research adopts a technological perspective to the use of big data and is mainly focused on mobile computing. Finally, the last cluster (in purple) is composed by the following three scholars: Akhtar, Rao-Nicholson, and Khan, with three papers and 64 citations each. These researchers share the authorship of multiple papers in the management field, some of which are related to Big Data analytics and sustainability in emerging markets. It should be noted that the last two clusters offer new perspectives on BD&S, not addressed before.

#### 4.7. Country and University Co-Author Analysis

The main purpose of co-authorship analysis is to reveal the structure of collaboration between institutions and countries at the researcher level and thus visualize the behavior of research teams [87]. The nodes reveal the influence of countries/institutions, while the thickness and distance among them show the degree of collaboration.

First, the collaboration among countries is analyzed. Figure 7 indicates that USA (142 documents, 2411 citations), China (140 documents, 2230 citations), England (101 documents, 1593 citations), India (62 documents, 1026 citations), Australia (55 documents, 1188 citations), Italy (53 documents, 602 citations), Spain (43 documents, 363 citations), France (41 documents, 695 citations), Norway (31 documents, 552 citations), and South Korea (26 documents, 194 citations) have significant influence on the BD&S literature (the software separates England, Scotland, and Wales instead of considering UK as a whole), according to the amount of publications. In the top 10 ranking, per number of citations, are also the U. Arab Emirates (6 documents, 530 citations) and Germany (25 documents, 489 citations).



**Figure 7.** Countries co-authorship network of BD&S: 44 countries of the 81 nations, which meet the threshold of a minimum number of papers per country of 5.

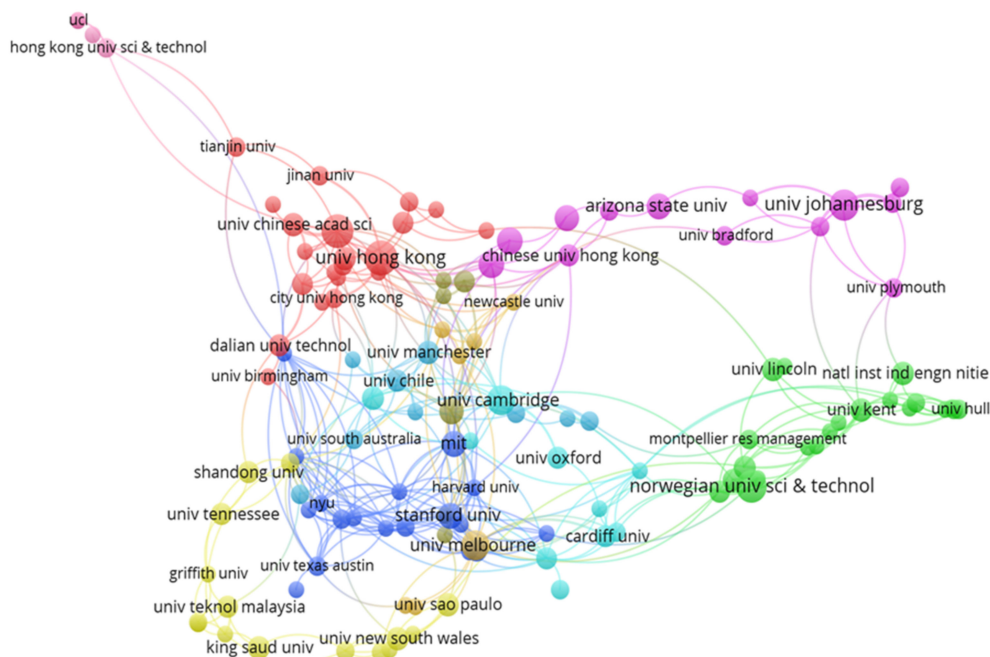
The conducted analysis reveals the existence of six clusters. The largest one (in red) includes 16 countries, which are basically European Union member countries (e.g., Italy, France, Spain, Norway). The second most important cluster is displayed in green and encompasses 11 countries: the USA, Australia, China, India, South Korea, and the U. Arab Emirates. The third cluster, in dark blue, contains 9 elements, among which is Malaysia (20 documents, 301 citations) together with other Asian countries, two European ones, and Mexico. The fourth cluster, in yellow, includes only Brazil, Chile, and Japan. The fifth one also contains 3 elements: Canada and, at a longer distance, Belgium and Romania. Finally, the light blue cluster includes only England and Taiwan.

Lastly, the most prominent institutions in BD&S research are analyzed (Figure 8). Judging again by the number of publications, the main universities are Norwegian University of Science and Technology (29 documents, 552 citations), University of Hong Kong (13 documents, 116 citations), Chinese Academy of Sciences (18 documents, 210 citations), University of Johannesburg (10 documents, 72 citations), University of Melbourne (9 documents, 156 citations), University of Cambridge (9 documents, 41 citations), Hong Kong Polytechnic University (8 documents, 220 citations), Shenzhen University (8 documents, 77 citations), Plymouth University (7 documents, 361 citations), University of Illinois (7 documents, 119 citations), Tsinghua University (7 documents, 188 citations), Arizona State University (7 documents, 61 citations), Massachusetts Institute of Technology (MIT) (7 documents, 55 citations), and Stanford University (7 documents, 24 citations).

Other relevant universities in the ranking of citations, with more than 200 citations are Beihang University (3 documents, 344 citations), University of Chile (5 documents, 326 citations), National Institute of Industrial Engineering (NITIE) (6 documents, 286 citations), and California State University (3 documents, 247 citations). While there are 148 institutions with more than 3 documents, only 124 are connected in 10 clusters. In the first cluster, with 21 items (in red), the University of Hong Kong and the Chinese Academy of Sciences are two of the top three institutions as per number of publications. The second cluster also contains 21 items and identifies the Norwegian University of Science and Technology as the most prolific one. The cluster also includes a large representation of French universities, as well as NITIE, and California State University. The dark blue cluster comprises 18 institutions (mainly located in the USA), among which are MIT and Stanford University. The fourth cluster, in yellow, contains 15 institutions and is anchored by the University of Tennessee. Furthermore, the cluster includes a relevant number of Australian universities.



The fifth group, colored in purple, comprises 13 items, among which are the leading universities in the publications ranking such as the University of Johannesburg, Hong Kong Polytechnic University, Shenzhen University, University of Plymouth, Tsinghua University, and Arizona State University. The most relevant institutions in the sixth network (in light blue) are the University of Cambridge and the National University of Singapore, as well as other English institutions. The seventh cluster, with 9 items, located in the middle of the figure, is anchored by the University of Chile and the University of Manchester. The eighth cluster, in light brown, contains seven institutions, with the University of Melbourne being the most prolific one. The ninth cluster includes the University of Illinois as its most relevant representative of this group of six research centers. The last cluster (pink) contains only four items.



**Figure 8.** Institution's co-authorship network of BD&S: 148 organizations of the 1174 meet the threshold of a minimum number of documents of 3.

## 5. Discussion

The present research has investigated the existing body of literature on the intersection between Big Data and Sustainability. Given the lack of a systematic understanding of this field of knowledge, the study has contributed to the literature by carrying out a bibliometric analysis and visualizing the most relevant pieces of research on BD&S. The conducted literature review on the concepts of “Big Data” and “sustainability” in the beginning of the paper also helps to improve the current understanding of the topic.

The results reveal the interdisciplinary nature of the area, integrating sustainability, technological and management perspectives. However, it should be noted that the literature on this topic is still in its infancy, and while it is currently expanding, most of the studies are conceptual and theoretical in nature. Furthermore, the extant advances are partial and limited, thus failing to integrate the developments in the main scientific fields.

While the research addressing Big Data and Sustainability has only started a decade ago, its growth has been significant in recent years, considering that the number of published articles has increased from 5 in 2014 to more than 300 in 2020. It should also be highlighted that 2020 was a particularly fruitful in terms of publications on BD&S since more than 40% of all the papers were published this year. This exponential growth in the number of publications is reflected by the impact of the some of the studies, with the Tao

et al. [78] article receiving more than 300 citations and Bibri and Krogstie [79] being cited more than 250 times in WoS, despite both been published less than 5 years ago.

Another relevant finding of the conducted bibliometric analysis is that the main academic fields investigating on BD&S are Green Sustainable Science Technology, Environmental Sciences, and Environmental Studies. The leading journal in terms of number of publications is Sustainability, followed, at a great distance, by the Journal of Cleaner Production and Sustainable Cities and Society. However, the existing articles are spread over 370 journals, indicating their expansion across a wide range of disciplines.

Table 5 summarizes the streams of research and main findings of the different bibliometric analyses.

**Table 5.** Main results of the bibliometric analyses.

Analysis	Main Clusters (Main Streams)	N° of Items
Co-occurrence network of "all keywords"	Sustainability	21
	Big data	18
	Management	14
	Framework	14
	Industry 4.0	6
Reference co-citation analysis	Big Data management	24
	Sustainability issues with geographical scopes (cities, urbanism)	14
	Sustainable manufacturing in industry 4.0 (technological perspective)	9
Journal co-citation analysis	Sustainability and Empirical sciences (Sustainability of cities)	34
	Management (Information System, decisions and Operations)	29
	Production Management	16
Authors' co-citation analysis	New technologies in energy	9
	Big Data (logistics and supply chain management)	35
	Smart sustainable cities and smart urbanism	29
	Sustainable manufacturing in industry 4.0	25
	Mathematical and engineering perspective	3
Bibliographic coupling of authors	Industry 4.0	18
	Smart sustainable cities	18
	Big Data analytics (logistics and supply chain management)	6
	Technological use of big data (mobile computing)	5
	Management (big data analytics and sustainability in emerging markets)	3
Country co-author analysis	European countries	16
	USA-China-India-Australia	11
	Malaysian and Asian Countries	9
	Brazil-Chile-Japan	3
	Canada-Belgium-Romania	3
	England-Taiwan	2
	U.Hong Kong-Chinese A.Sc.	21
	Norwegian U.Sc.-French universities	21
	USA Universities (MIT, Stanford)	18
	USA-Australian Universities (Tennessee)	15
University co-author analysis	U.Johannesbourg-Hong Kong P.U.	13
	English institutions (U. Cambridge)-N.U.	10
	Singapore	
	U.Chile-U.Manchester	9
	U.Melbourne	7
	U.Illinois	6
	Hong Kong U.Sc.T.	4

As for the keyword co-occurrence, the most frequent keywords are "big data", "sustainability", "management", "framework", and "challenges", which is indicative of the environmental, technological, and managerial perspectives adopted to study this topic. Industry 4.0 and supply chain management appears as another cluster independent from the management one. The cluster labelled "framework", as well as the frequently used

keywords “challenges”, “future”, and “model” suggest that the literature on BD&S is in a conceptualization phase.

The co-citation analysis reveals that the papers by Wang et al. [82], McAfee et al. [83], and Kitchin [84] are the most cited in the research sample. The first two focus on Big Data management and represent a cluster that addresses the company management of these technological developments for sustainability purposes. The third anchors the second cluster, which adopts an environmental perspective, is associated with geographical area management, mainly looking at the development of smart cities. The third group is closely related to the first one, while the last one deals with Industry 4.0 and the Internet of Things applying a more technological approach.

The study of journal co-citations also uncovers four groups. The first one, anchored by Sustainability, the third most cited source, is centered on the sustainability of smart cities. The second one, in contrast, has a strong management focus and mainly investigates information systems for decision-making. This cluster is anchored by the International Journal of Production Economics, the second most cited source. The Journal of Clean Production, the most cited source, is the most relevant one in the third cluster, which also has a predominantly management focus, although it mainly addresses production and operations management. While the fourth group is more residual, it reveals a nascent research area on technological innovations related to renewable development in the energy field.

The analysis of co-citations of authors also identifies four clusters. The first one, anchored by Dubey, the second most cited author, investigates aspects associated with Big Data, in general, and Big Data Analytics in the logistics and supply chain management field. The second largest cluster, which includes Bibri and Kitchin (the first and third most cited authors), also focuses on smart sustainable cities and smart urbanism. The third group addresses sustainable manufacturing in industry 4.0. Finally, the fourth network is centered on digital technology and smart factory industry 4.0, thus representing a new research line that adopts mathematical and engineering perspectives.

The bibliographic coupling of authors indicates the existence of 6 groups of authors, the largest one of which is anchored by the fourth most cited author, Bag, and focuses on industry 4.0. The second most important cluster is dominated by Bibri and Krogstie and addresses smart sustainable cities, as mentioned above. The third group, anchored by Gunasekaran, deals with Big Data analytics in logistics and supply chain management. The fourth cluster looks at Big Data through a technological lens, with a particular emphasis on mobile computing. The fifth group opens a new perspective on Big Data analytics and sustainability in emerging markets.

The co-authorship analysis of BD&S shows the leadership of the USA, China, and England. However, the most important cluster represents mainly European Union countries. The second group includes the USA and China, together with India and Australia. The third group represents mainly Asian countries, while the fourth one comprises South American ones. As for the rest of the clusters, it should be noted that they include geographically dispersed locations. Regarding the main institutions contributing to the knowledge on BD&S, the Norwegian University of Science and Technology, the University of Hong Kong and the Chinese Academy of Sciences are revealed as the most prolific centers. The first cluster is anchored by multiple Chinese institutions, while the most relevant in the second one is the Norwegian University of Science and Technology. The third cluster subsumes mainly US institutions, while the fourth one comprises Australian universities. Authors should discuss the results and how they can be interpreted from the perspective of previous studies and of the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted.

## 6. Conclusions

This study has analyzed the state of the art, and trends, in the BD&S literature. The results have shown the existence of various theoretical research areas, which have

approached the topic from heterogeneous perspectives. Overall, the analysis has revealed the existence of four main approaches to the study of sustainability and big data: an information systems management perspective; an environmental approach; an operations management perspective; the technological stance related to the so-called Industry 4.0 and the Internet of Things. In addition, a few minor approaches, which may evolve in new research trends, were unveiled: mobile computing, mathematics, engineering, as well as specific developments in the field of energy. The study emphasizes the relevance of these emerging areas and trends, which are still in a predominantly theoretical and conceptual stage of development. The findings are important for both academic scholars and practitioners, since they anticipate not only new lines of research, but also key aspects that have to be considered in the practical implementation of BD&S.

In terms of practical implications, the study uncovers opportunities for new business models, improved decision-making, as well as innovations in the operations management and the internal and external logistics chain. In addition, the identified developments in the literature on smart cities also offer implications for the decision-making of governmental institutions. However, the study also visualizes the lack of Big Data application in business functions, such as marketing and the company's internal value chain, among others. The application of big data is also scarce across most production sectors, except for the few studies conducted in the energetic field. In this regard, the lack of BD&S studies in the healthcare and other services sectors, which have a high impact on today's economies is particularly notable. In addition, the developments observed in the business field should be transferred to the decision-making of public institutions beyond issues such as sustainable urban planning and renewable energy management.

As for the theoretical relevance of this article, the conducted bibliometric study opens new research avenues on BD&S. The paper visualizes the embryonic state of the art of the literature, which is still in a conceptual stage, and thus opens diverse methodological applications, as observed by Garrigos et al. [2,71] in other bibliometric analyses. Furthermore, as previously indicated, the application of the BD&S is still lacking in multiple sectors. In addition, academic research going beyond the technological and management approaches is still scarce. In this regard, and although some developments in the fields of mathematics, geography and engineering have been revealed, the rest of research areas have not addressed the topic yet. These would include health care, biology, physics, and environmental sciences. As for the social sciences, the sustainable development of big data should be approached by sociology, ethics, political science, education sciences, and economics.

In this regard, the present study can uncover new research trends, and help researchers to discover important gaps. Future investigations should consider the aspects pointed out by Garrigos et al. [2,71]: issues related to trends in BD&S in classical research fields, the application to new areas and sectors where this literature has not yet developed, and the consideration of BD&S when analyzing timely topics. In this sense, the findings regarding the most influential and most cited authors, sources, and papers, as well as the listing of the most frequent keywords, can help scholars to identify the most relevant research questions and the hot topics that attract the current research interest.

Furthermore, both academic studies and practical applications should consider the link between business and technological developments, since the management and the technological side of BD&S are the two main areas that have been uncovered by the conducted bibliometric analysis (apart from sustainable management of smart cities). Hence, future research efforts and practical applications should be cognizant of the evolution of the Web 3.0 or Semantic Web and the conceptual models pointed out by Garrigos et al. [1] where Artificial Intelligence, as well as the Ubiquitous and Pervasive Web associated with the Internet of Things play a crucial role. In addition, the evolution of Big Data itself should also be closely followed, including not only the numerical and alphabetical data from organizations' internal warehouses, but also the management and mining of data from social networks, external websites, smartphones, data in the cloud, which is increasingly decentralized, and associated with images, tastes, and other senses. In addition, the new

developments should consider the new types of companies, which are essentially virtual organizations and outsourced networks, where crowdsourcing plays a key role [62].

Finally, future research may attempt to overcome the limitations of this study. Further investigations can apply new methodologies, incorporate qualitative data, or adopt a more focused approach to BD&S and thus, address the limitations of the conducted bibliometric and visualization analyses. Some fruitful areas of research would be to complement the study with a more extensive sample from additional databases; to consider other types of documents such as conference, proceedings, and professional papers; to analyze documents published in other languages or to focus on a specific geographical area. Future studies could also develop the main trends identified by this research, further develop, and complement to the conducted bibliometric analyses, as well as implement new techniques presenting relevant methodological innovations.

**Author Contributions:** Conceptualization, F.G.-S. and S.S.-B.; methodology, F.G.-S.; software, F.G.-S.; validation, Y.N., S.S.-B.; formal analysis, F.G.-S. and S.S.-B.; resources, F.G.-S. and S.S.-B.; data curation, F.G.-S.; writing—original draft preparation, Y.N. and S.S.-B.; writing—review and editing, Y.N. and D.B.; project administration, D.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data is available in the Web of Science Core Collection database.

**Conflicts of Interest:** The authors declare no conflict of interests.

## References

- Garrigos, F.; Lapedra, R.; Barberá, T. Social networks and Web 3.0: Their impact on the management and marketing of organizations. *Manag. Decis.* **2012**, *50*, 1880–1889. [\[CrossRef\]](#)
- Garrigos-Simon, F.J.; Narangajavana-Kaosiri, Y.; Lengua-Lengua, I. Tourism and Sustainability: A Bibliometric and Visualization Analysis. *Sustainability* **2018**, *10*, 1976. [\[CrossRef\]](#)
- Broadus, R. Toward a definition of “bibliometrics”. *Scientometrics* **1987**, *12*, 373–379. [\[CrossRef\]](#)
- Ahmad, I.; Ahmed, G.; Shah, S.A.A.; Ahmed, E. A decade of big data literature: Analysis of trends in light of bibliometrics. *J. Supercomput.* **2018**, *76*, 3555–3571. [\[CrossRef\]](#)
- Gupta, D.; Rani, R. A study of big data evolution and research challenges. *J. Inf. Sci.* **2019**, *45*, 322–340. [\[CrossRef\]](#)
- Hu, J.; Zhang, Y. Discovering the interdisciplinary nature of Big Data research through social network analysis and visualization. *Scientometrics* **2017**, *112*, 91–109. [\[CrossRef\]](#)
- Hu, F.; Liu, W.; Tsai, S.-B.; Gao, J.; Bin, N.; Chen, Q. An Empirical Study on Visualizing the Intellectual Structure and Hotspots of Big Data Research from a Sustainable Perspective. *Sustainability* **2018**, *10*, 667. [\[CrossRef\]](#)
- Liu, X.; Sun, R.; Wang, S.; Wu, Y.J. The research landscape of big data: A bibliometric analysis. *Libr. Hi Tech* **2019**, *38*, 367–384. [\[CrossRef\]](#)
- Mazieri, M.; Soares, E. Conceptualization and theorization of the Big Data. *Int. J. Innov.* **2016**, *4*, 23–41. [\[CrossRef\]](#)
- Peng, Y.; Shi, J.; Fantinato, M.; Chen, J. A study on the author collaboration network in big data. *Inf. Syst. Front.* **2017**, *19*, 1329–1342. [\[CrossRef\]](#)
- Saheb, T.; Saheb, T. Understanding the development trends of big data technologies: An analysis of patents and the cited scholarly works. *J. Big Data* **2020**, *7*, 1–26. [\[CrossRef\]](#)
- Aboelmaged, M.; Mouakket, S. Influencing models and determinants in big data analytics research: A bibliometric analysis. *Inf. Process. Manag.* **2020**, *57*, 102234. [\[CrossRef\]](#)
- Inamdar, Z.; Raut, R.; Narwane, V.S.; Gardas, B.; Narkhede, B.; Sagnak, M. A systematic literature review with bibliometric analysis of big data analytics adoption from period 2014 to 2018. *J. Enterp. Inf. Manag.* **2020**, *34*, 101–139. [\[CrossRef\]](#)
- Zhang, Y.; Huang, Y.; Porter, A.L.; Zhang, G.; Lu, J. Discovering and forecasting interactions in big data research: A learning-enhanced bibliometric study. *Technol. Forecast. Soc. Chang.* **2019**, *146*, 795–807. [\[CrossRef\]](#)
- Liao, H.; Tang, M.; Luo, L.; Li, C.; Chiclana, F.; Zeng, X.-J. A Bibliometric Analysis and Visualization of Medical Big Data Research. *Sustainability* **2018**, *10*, 166. [\[CrossRef\]](#)
- Galetsi, P.; Katsaliaki, K. Big data analytics in health: An overview and bibliometric study of research activity. *Health Inf. Libr. J.* **2020**, *37*, 5–25. [\[CrossRef\]](#)
- Gu, D.; Li, J.; Li, X.; Liang, C. Visualizing the knowledge structure and evolution of big data research in healthcare informatics. *Int. J. Med. Inform.* **2017**, *98*, 22–32. [\[CrossRef\]](#)
- Hashem, I.A.T.; Anuar, N.B.; Gani, A.; Yaqoob, I.; Xia, F.; Khan, S.U. MapReduce: Review and open challenges. *Scientometrics* **2016**, *109*, 389–422. [\[CrossRef\]](#)

19. Belmonte, J.L.; Segura-Robles, A.; Moreno-Guerrero, A.-J.; Parra-González, M.E. Machine Learning and Big Data in the Impact Literature. A Bibliometric Review with Scientific Mapping in Web of Science. *Symmetry* **2020**, *12*, 495. [[CrossRef](#)]
20. Khanra, S.; Dhir, A.; Mäntymäki, M. Big data analytics and enterprises: A bibliometric synthesis of the literature. *Enterp. Inf. Syst.* **2020**, *14*, 737–768. [[CrossRef](#)]
21. Aykroyd, R.G.; Leiva, V.; Ruggeri, F. Recent developments of control charts, identification of big data sources and future trends of current research. *Technol. Forecast. Soc. Chang.* **2019**, *144*, 221–232. [[CrossRef](#)]
22. Wamba, S.F.; Mishra, D. Big data integration with business processes: A literature review. *Bus. Process Manag. J.* **2017**, *23*, 477–492. [[CrossRef](#)]
23. Arunachalam, D.; Kumar, N.; Kawalek, J.P. Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. *Transp. Res. Part. E Logist. Transp. Rev.* **2018**, *114*, 416–436. [[CrossRef](#)]
24. Mishra, D.; Gunasekaran, A.; Papadopoulos, T.; Childe, S.J. Big Data and supply chain management: A review and bibliometric analysis. *Ann. Oper. Res.* **2018**, *270*, 313–336. [[CrossRef](#)]
25. López-Robles, J.; Otegi-Olaso, J.; Gómez, I.P.; Cobo, M.J. 30 years of intelligence models in management and business: A bibliometric review. *Int. J. Inf. Manag.* **2019**, *48*, 22–38. [[CrossRef](#)]
26. Chen, H.; Chiang, R.H.L.; Storey, V.C. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Q.* **2012**, *36*, 1165. [[CrossRef](#)]
27. Liang, T.-P.; Liu, Y.-H. Research Landscape of Business Intelligence and Big Data analytics: A bibliometrics study. *Expert Syst. Appl.* **2018**, *111*, 2–10. [[CrossRef](#)]
28. Batistič, S.; Van Der Laken, P. History, Evolution and Future of Big Data and Analytics: A Bibliometric Analysis of Its Relationship to Performance in Organizations. *Br. J. Manag.* **2019**, *30*, 229–251. [[CrossRef](#)]
29. Rialti, R.; Marzi, G.; Ciappei, C.; Busso, D. Big data and dynamic capabilities: A bibliometric analysis and systematic literature review. *Manag. Decis.* **2019**, *57*, 2052–2068. [[CrossRef](#)]
30. Vanhala, M.; Lu, C.; Peltonen, J.; Sundqvist, S.; Nummenmaa, J.; Järvelin, K. The usage of large data sets in online consumer behaviour: A bibliometric and computational text-mining-driven analysis of previous research. *J. Bus. Res.* **2020**, *106*, 46–59. [[CrossRef](#)]
31. Song, M.; Kim, S.; Zhang, G.; Ding, Y.; Chambers, T. Productivity and influence in bioinformatics: A bibliometric analysis using PubMed central. *J. Assoc. Inf. Sci. Technol.* **2014**, *65*, 352–371. [[CrossRef](#)]
32. Firdaus, A.; Ab Razak, M.F.; Feizollah, A.; Hashem, I.A.T.; Hazim, M.; Anuar, N.B. The rise of “blockchain”: Bibliometric analysis of blockchain study. *Scientometrics* **2019**, *120*, 1289–1331. [[CrossRef](#)]
33. Ruiz-Rosero, J.; Ramirez-Gonzalez, G.; Williams, J.M.; Liu, H.; Khanna, R.; Pisharody, G. Internet of Things: A Scientometric Review. *Symmetry* **2017**, *9*, 301. [[CrossRef](#)]
34. Zhang, Y.; Hua, W.; Yuan, S. Mapping the scientific research on open data: A bibliometric review. *Learn. Publ.* **2018**, *31*, 95–106. [[CrossRef](#)]
35. Ivanov, D.; Tang, C.S.; Dolgui, A.; Battini, D.; Das, A. Researchers’ perspectives on Industry 4.0: Multi-disciplinary analysis and opportunities for operations management. *Int. J. Prod. Res.* **2021**, *59*, 2055–2078. [[CrossRef](#)]
36. Kipper, L.M.; Furstenau, L.B.; Hoppe, D.; Frozza, R.; Iepsen, S. Scopus scientific mapping production in industry 4.0 (2011–2018): A bibliometric analysis. *Int. J. Prod. Res.* **2019**, *58*, 1605–1627. [[CrossRef](#)]
37. Nazarov, D.; Klarin, A. Taxonomy of Industry 4.0 research: Mapping scholarship and industry insights. *Syst. Res. Behav. Sci.* **2020**, *37*, 535–556. [[CrossRef](#)]
38. Da Costa, M.B.; dos Santos, L.M.A.L.; Schaefer, J.L.; Baierle, I.C.; Nara, E.O.B. Industry 4.0 technologies basic network identification. *Scientometrics* **2019**, *121*, 977–994. [[CrossRef](#)]
39. Kulakli, A.; Osmanaj, V. Global Research on Big Data in Relation with Artificial Intelligence (A Bibliometric Study: 2008–2019). *Int. J. Online Biomed. Eng. (ijOE)* **2020**, *16*, 31–46. [[CrossRef](#)]
40. Raban, D.R.; Gordon, A. The evolution of data science and big data research: A bibliometric analysis. *Scientometrics* **2020**, *122*, 1563–1581. [[CrossRef](#)]
41. Nobre, G.C.; Tavares, E. Scientific literature analysis on big data and internet of things applications on circular economy: A bibliometric study. *Scientometrics* **2017**, *111*, 463–492. [[CrossRef](#)]
42. Gobbo, J.A.; Busso, C.M.; Gobbo, S.C.O.; Carreão, H. Making the links among environmental protection, process safety, and industry 4.0. *Process Saf. Environ. Prot.* **2018**, *117*, 372–382. [[CrossRef](#)]
43. Felsberger, A.; Reiner, G. Sustainable Industry 4.0 in Production and Operations Management: A Systematic Literature Review. *Sustainability* **2020**, *12*, 7982. [[CrossRef](#)]
44. Della Corte, V.; Del Gaudio, G.; Sepe, F.; Sciarelli, F. Sustainable Tourism in the Open Innovation Realm: A Bibliometric Analysis. *Sustainability* **2019**, *11*, 6114. [[CrossRef](#)]
45. Sharma, R.; Jabbour, C.J.C.; de Sousa Jabbour, A.B.L. Sustainable manufacturing and industry 4.0: What we know and what we don’t. *J. Enterp. Inf. Manag.* **2020**, *34*, 230–266. [[CrossRef](#)]
46. Zhao, L.; Tang, Z.-Y.; Zou, X. Mapping the Knowledge Domain of Smart-City Research: A Bibliometric and Scientometric Analysis. *Sustainability* **2019**, *11*, 6648. [[CrossRef](#)]

47. Kong, L.; Liu, Z.; Wu, J. A systematic review of big data-based urban sustainability research: State-of-the-science and future directions. *J. Clean. Prod.* **2020**, *273*, 123142. [[CrossRef](#)]
48. Chalmeta, R.; Santos-Deleón, N.J. Sustainable Supply Chain in the Era of Industry 4.0 and Big Data: A Systematic Analysis of Literature and Research. *Sustainability* **2020**, *12*, 4108. [[CrossRef](#)]
49. Zhang, X.; Yu, Y.; Zhang, N. Sustainable supply chain management under big data: A bibliometric analysis. *J. Enterp. Inf. Manag.* **2020**, *34*, 427–445. [[CrossRef](#)]
50. Cappa, F.; Oriani, R.; Peruffo, E.; McCarthy, I. Big Data for Creating and Capturing Value in the Digitalized Environment: Unpacking the Effects of Volume, Variety, and Veracity on Firm Performance. *J. Prod. Innov. Manag.* **2021**, *38*, 49–67. [[CrossRef](#)]
51. Sanchez-Planelles, J.; Segarra-Oña, M.; Peiro-Signes, A. Building a Theoretical Framework for Corporate Sustainability. *Sustainability* **2021**, *13*, 273. [[CrossRef](#)]
52. Jin, X.; Wah, B.W.; Cheng, X.; Wang, Y. Significance and Challenges of Big Data Research. *Big Data Res.* **2015**, *2*, 59–64. [[CrossRef](#)]
53. Guo, W. Using Semantic Web technologies for ubiquitous computing. In Proceedings of the 2008 First IEEE International Conference on Ubi-Media Computing, Lanzhou, China, 31 July–1 August 2008; pp. 377–381.
54. Singh, S.; Puradkar, S.; Lee, Y. Ubiquitous computing: Connecting Pervasive computing through Semantic Web. *Inf. Syst. e-Business Manag.* **2006**, *4*, 421–439. [[CrossRef](#)]
55. Alfouzan, H.I. Big Data In Business. *Int. J. Sci. Eng. Res.* **2015**, *6*, 1351–1352.
56. Alsgaier, H.; Akour, M.; Shehabat, I.; Aldiabat, S. The Importance of Big Data Analytics in Business: A Case Study. *Am. J. Softw. Eng. Appl.* **2017**, *6*, 111. [[CrossRef](#)]
57. Franco, S. The influence of the external and internal environments of multinational enterprises on the sustainability commitment of their subsidiaries: A cluster analysis. *J. Clean. Prod.* **2021**, *297*, 126654. [[CrossRef](#)]
58. Parviainen, P.; Tihinen, M.; Kääriäinen, J.; Teppola, S. Tackling the digitalization challenge: How to benefit from digitalization in practice. *Int. J. Inf. Syst. Project Manag.* **2017**, *5*, 63–77.
59. Galdon-Salvador, J.L.; Garrigos-Simon, F.J.; Gil-Pechuan, I. Improving hotel industry processes through crowdsourcing techniques. In *Open Tourism: Open Innovation, Crowdsourcing and Co-Creation Challenging the Tourism Industry*; Egger, R., Gula, I., Walcher, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 95–107.
60. Garrigos-Simon, F.J.; Narangajavana, Y.; Galdón-Salvador, J.L. Crowdsourcing as a competitive advantage for new business models. In *Strategies in E-Business*; Gil-Pechuán, I., Palacios-Marqués, D., Peris Ortiz, M.P., Eds.; Springer: Boston, MA, USA, 2014; pp. 29–37.
61. Garrigos-Simon, F.J.; Narangajavana, Y. From Crowdsourcing to the Use of Masscapital. The Common Perspective of the Success of Apple, Facebook, Google, Lego, TripAdvisor, and Zara. In *Advances in Crowdsourcing*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 1–13.
62. Garrigos-Simon, F.J.; Gil-Pechuán, I.; Estelles-Miguel, S. (Eds.) *Advances in Crowdsourcing*; Springer: Cham, Switzerland, 2015; pp. 1–183.
63. Garrigos-Simon, F.J.G.; Llorente, R.; Morant, M.; Narangajavana, Y. Pervasive information gathering and data mining for efficient business administration. *J. Vacat. Mark.* **2016**, *22*, 295–306. [[CrossRef](#)]
64. Zhu, X.; Yang, Y. Big Data Analytics for Improving Financial Performance and Sustainability. *J. Syst. Sci. Inf.* **2021**, *9*, 175–191. [[CrossRef](#)]
65. Dubey, R.; Gunasekaran, A.; Childe, S.J.; Papadopoulos, T.; Luo, Z.; Wamba, S.F.; Roubaud, D. Can big data and predictive analytics improve social and environmental sustainability? *Technol. Forecast. Soc. Chang.* **2019**, *144*, 534–545. [[CrossRef](#)]
66. Duvnjak, K.; Gregorić, M.; Gorše, M. Sustainable development—an artificial intelligence approach. *Manag. Res. Pract.* **2020**, *12*, 18–28.
67. Visconti, R.M.; Morea, D. Big Data for the Sustainability of Healthcare Project Financing. *Sustainability* **2019**, *11*, 3748. [[CrossRef](#)]
68. Runting, R.K.; Phinn, S.; Xie, Z.; Venter, O.; Watson, J.E.M. Opportunities for big data in conservation and sustainability. *Nat. Commun.* **2020**, *11*, 2003. [[CrossRef](#)] [[PubMed](#)]
69. Merigó, J.M.; Yang, J.B. Accounting Research: A Bibliometric Analysis. *Aust. Account. Rev.* **2017**, *27*, 71–100. [[CrossRef](#)]
70. Delgado López-Cózar, E.; Robinson-García, N.; Torres-Salinas, D. The Google scholar experiment: How to index false papers and manipulate bibliometric indicators. *J. Assoc. Inf. Sci. Technol.* **2014**, *65*, 446–454. [[CrossRef](#)]
71. Garrigos-Simon, F.J.; Narangajavana-Kaosiri, Y.; Narangajavana, Y. Quality in Tourism Literature: A Bibliometric Review. *Sustainability* **2019**, *11*, 3859. [[CrossRef](#)]
72. Blanco-Mesa, F.R.B.; Merigó, J.M.; Gil Lafuente, A.M. Fuzzy decision making: A bibliometric-based review. *J. Intell. Fuzzy Syst.* **2017**, *32*, 2033–2050. [[CrossRef](#)]
73. Van Eck, N.J.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [[CrossRef](#)]
74. Small, H. Co-citation in the scientific literature: A new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* **1973**, *24*, 265–269. [[CrossRef](#)]
75. Kessler, M.M. Bibliographic coupling between scientific papers. *Am. Doc.* **1963**, *14*, 10–25. [[CrossRef](#)]
76. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572. [[CrossRef](#)]

77. Toole, J.L.; Eagle, N.; Plotkin, J.B. Spatiotemporal correlations in criminal offense records. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1–18. [[CrossRef](#)]
78. Tao, F.; Cheng, J.; Qi, Q.; Zhang, M.; Zhang, H.; Sui, F. Digital twin-driven product design, manufacturing and service with big data. *Int. J. Adv. Manuf. Technol.* **2018**, *94*, 3563–3576. [[CrossRef](#)]
79. Bibri, S.E.; Krogstie, J. Smart sustainable cities of the future: An extensive interdisciplinary literature review. *Sustain. Cities Soc.* **2017**, *31*, 183–212. [[CrossRef](#)]
80. Al Nuaimi, E.; Al Neyadi, H.; Mohamed, N.; Al-Jaroodi, J. Applications of big data to smart cities. *J. Internet Serv. Appl.* **2015**, *6*, 25. [[CrossRef](#)]
81. Wang, Y.; Chen, Q.; Hong, T.; Kang, C. Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. *IEEE Trans. Smart Grid* **2019**, *10*, 3125–3148. [[CrossRef](#)]
82. Wang, G.; Gunasekaran, A.; Ngai, E.W.; Papadopoulos, T. Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *Int. J. Prod. Econ.* **2016**, *176*, 98–110. [[CrossRef](#)]
83. McAfee, A.; Brynjolfsson, E.; Davenport, T.H.; Patil, D.J.; Barton, D. Big data: The management revolution. *Harv. Bus. Rev.* **2012**, *90*, 60–68. [[PubMed](#)]
84. Kitchin, R. The real-time city? Big data and smart urbanism. *GeoJournal* **2014**, *79*, 1–14. [[CrossRef](#)]
85. Gunasekaran, A.; Papadopoulos, T.; Dubey, R.; Wamba, S.F.; Childe, S.J.; Hazen, B.; Akter, S. Big data and predictive analytics for supply chain and organizational performance. *J. Bus. Res.* **2017**, *70*, 308–317. [[CrossRef](#)]
86. Stock, T.; Seliger, G. Opportunities of Sustainable Manufacturing in Industry 4.0. *Procedia CIRP* **2016**, *40*, 536–541. [[CrossRef](#)]
87. Reyes-Gonzalez, L.; Gonzalez-Brambila, C.N.; Veloso, F. Using co-authorship and citation analysis to identify research groups: A new way to assess performance. *Scientometrics* **2016**, *108*, 1171–1191. [[CrossRef](#)]