


Review

Sustainable Supply Chain in the Era of Industry 4.0 and Big Data: A Systematic Analysis of Literature and Research

Ricardo Chalmeta *  and Nestor J. Santos-deLeón

Grupo de Investigación en Integración y Re-Ingeniería de sistemas (IRIS), Departamento de Lenguajes y Sistemas informáticos, Universitat Jaume I, 12006 Castellón, Spain; nsantosdeleon@gmail.com

* Correspondence: rchalmet@uji.es; Tel.: +34-964-728435

Received: 22 March 2020; Accepted: 13 May 2020; Published: 18 May 2020



Abstract: Supply chain sustainability (SCS) in the age of Industry 4.0 and Big Data is a growing area of research. However, there are no systematic and extensive studies that classify the different types of research and examine the general trends in this area of research. This paper reviews the literature on sustainability, Big Data, Industry 4.0 and supply chain management published since 2009 and provides a thorough insight into the field by using bibliometric and network analysis techniques. A total of 87 articles published in the past 10 years were evaluated and the top contributing authors, countries, and key research topics were identified. Furthermore, the most influential works based on citations and PageRank were obtained and compared. Finally, six research categories were proposed in which scholars could be encouraged to expand Big Data and Industry 4.0 research on SCS. This paper contributes to the literature on SCS in the age of Industry 4.0 by discussing the challenges facing current research but also, more importantly, by identifying and proposing these six research categories and future research directions.

Keywords: sustainability; supply chain; Big Data; Industry 4.0

1. Introduction

Sustainable development is defined as development that meets the needs of the present without compromising the ability of future generations to meet their own needs [1], and it has become a key strategic objective worldwide. Sustainable development requires the consideration and integration of economic, social, cultural, political, and ecological factors in decision-making, in an attempt to balance economic development, social development, and environmental protection [2]. As a result, the majority of large companies and an important number of small and medium-sized enterprises have been incorporating policies and actions aimed at improving their sustainability and the sustainability of their supply chains (SCs) [3]. This consideration of sustainability as an objective in supply chain management, in what is also known as Sustainable Supply Chain Management (SSCM), is due mainly to three factors: (1) the pressure of stakeholders (such as investors, shareholders, customers, and non-profits) to alleviate the enormous environmental impacts that are being generated (deterioration of the environment, scarcity of resources, increase in waste generated, increased pollution); (2) generating brand value that serves as a differentiating element against competitors; and (3) increasingly restrictive regulations [4].

However, the incorporation of sustainability into supply chain management faces a number of obstacles and difficulties. Pagell [5] states that present knowledge in SSCM is not sufficient to create truly sustainable supply chains and identifies problems with assumptions, norms, institutions, measures, and methods that future research needs to address. In addition to these shortcomings, Luthra and Mangla [6] add another: the technical issues related to the generation, processing,

and analysis of data that will allow greater effectiveness and efficiency in business processes, as well as performance control and support for supply chain decision-making.

This is where the concept of Industry 4.0 and the use of Big Data technology come into play. Industry 4.0 has become a buzzword that describes the trend towards digitalization and automation of manufacturing. It allows products, machines, components, individuals, and systems to create a smart network so that it can integrate cyber-physical systems to act quickly by linking information and physical memory to the smart network for faster and more effective service environments [7]. The contribution of Industry 4.0 to more sustainable industrial value creation will be remarkable in the future [8]. For example, it will allow efficient resource allocation, which includes water, energy, raw material, and other products, based on data that is collected in real time, resulting in new sustainable green practices. On the other hand, Big Data analytics are technologies and techniques used to analyse large-scale, complex data from various applications in order to acquire intelligence and to extract unknown, hidden, valid, and useful relationships, patterns and information [9]. Big Data can be very useful in enhancing supply chain sustainability as it can be used to identify the sustainability impact of the SC in the past, as well as to predict its future sustainability impact [10]. Therefore, the combined use of digitalization and automation of Industry 4.0 together with Big Data technology has great potential to overcome some of the problems affecting sustainable supply chain management [11].

However, development of the research on Big Data and Industry 4.0 to enhance supply chain sustainability is still limited, its final application in real cases of the business world is usually in the early stages [12], and hence practitioners have problems using Big Data and Industry 4.0 in supply chain sustainability [13]. This could limit its impact, since existing SSCM literature is primarily focused on other problems, such as building the definitions of SSCM; discussing the implementation of SSCM; proposing strategic decisions to be considered in SSCM; analysing SSCM governance mechanisms; and developing frameworks for SSCM [14]. Therefore, more research in the technological issues concerning the Industry 4.0 and Big Data solutions that support sustainable practices in the supply chain is required [6,15,16].

In order to cover the research gap between the growing and useful interest in the ways Industry 4.0 and Big Data can help to support sustainable supply chains, and the scarcity of systematic and extensive reviews of the recent research on this subject, this paper seeks to explore the status of research in the domains of Industry 4.0, Big Data, and sustainable supply chains, as well as to establish a categorical framework that brings together research conducted on the basis of relevant common points. The hypothesis is that a bibliographic analysis of current research can facilitate the advancement of future research in this area. In particular, the following research questions (RQs) are elaborated:

RQ1: Which are the top contributing authors, countries, and institutions in the field of Big Data and Industry 4.0 as applied to SSCM?

RQ2: Is it possible to define research categories on the basis of relevant common points?

RQ3: Which are the future research necessities in the field of Big Data and Industry 4.0 as applied to SSCM?

To answer the above research questions and to address Industry 4.0 and Big Data as means to support sustainable supply chains, this paper (1) reviews the literature on Big Data, Industry 4.0, and supply chain management published since 2009; (2) provides a thorough insight into the field by using bibliometric and network analysis techniques and by evaluating 87 articles published over the past 10 years, identifying top contributing authors, countries, and key research topics related to the field; (3) obtains and compares the most influential works based on citations and PageRank; (4) identifies and proposes six established and emerging research categories that would encourage scholars to expand research on Industry 4.0, Big Data, and SSCM; and (5) identifies the future research necessities in every category. The methodology we followed combines Rowley and Slack's proposal [17] and the snowball method [18] for item selection. To perform the bibliographic and network analyses, the paper by

Mishra et al. [19] was used as a model. Finally, for the identification of the categories, we followed the comparative method proposed by Collier [20].

This paper is organized as follows: Section 2 presents a review of the literature related to supply chain sustainability, Industry 4.0, and Big Data. Section 3 describes the software tools and the research methodology to perform the bibliographic and network analyses. Section 4 shows the findings of the bibliographic and network analyses. Finally, Section 5 discusses the findings and shows the conclusions, research limitations, and future work.

2. Literature Review

2.1. Sustainable Supply Chain Management

A Supply Chain can be defined as “a set of three or more entities (organizations or individuals) directly involved in the upstream and downstream flows of products, services, finances, and/or information from a source to a customer” [21]. The supply chain should be designed, managed, and coordinated as a single entity rather than separately as individual functions. Therefore, an adequate Supply Chain Management (SCM) must

- introduce the concept of continuous improvement and the need to be customer-oriented to the supply chain employees [22];
- optimize the two-directional flow of information, goods, technology, knowledge, human resources, and services among all the components of the chain [23];
- achieve both specific and common goals to improve long-term performance, for each company and for the supply chain as a whole [24].

Nowadays, one of the main challenges in SCM is the integration of sustainability principles into the supply chain, considering a multidimensional (the economic, environmental, and social impact) and multiscale approach (institutional, geographical, and temporal) [25]. The objective is to adopt social and environmental practices in the supply chain, in alignment with diverse stakeholder expectations, to mitigate sustainability-related risks [26]. This will result in the emergence of a new stream inside SCM—the sustainable supply chain management (SSCM). SSCM can be defined as “the designing, organizing, coordinating, and controlling of supply chains to become truly sustainable with the minimum expectation of a truly sustainable supply chain being to maintain economic viability, while doing no harm to social or environmental systems” [5].

There is a growing movement towards adopting social and environmental practices in the supply chain [27]. Many papers discuss the drivers and enablers for organizations implementing SSCM [14]. After an analysis of literature, Zimon et al. [28] summarize these drivers as Management commitment, Organisational involvement, Supportive culture, Productivity improvement, Waste elimination, Competitive opportunity, Business social compliance, Environmental regulation compliance, Green product requirement, Reverse logistics requirement, Customer and supplier involvement, Regulatory pressure, Institutional pressures, International environmental regulation, Competition, Reputation, and Social responsibility. Therefore, SSCM becomes a synergistic and dynamic part of corporate competitive advantage; and understanding how SSCM contributes to the creation of value is an important part of management [29].

SSCM involves different organizational, human, and technological elements, such as: (1) the involvement of all members of the supply chain in defining a new mission, vision, values, policies, objectives, and strategies of the SC that include the sustainability approach; (2) the redesign of business processes and the creation of new ones such as reverse logistics or close loop supply chains to incorporate sustainable practices; (3) training and generation of new skills among human resources; and (4) the support of information and communications technologies to optimize business processes and decision-making. This makes SSCM a complex task that must overcome different difficulties such as the diversity of products and services provided by the supply chain [30], the geographical dispersion

of the members of the supply chain with different legal regulations [31], the necessity of measuring the impacts of the supply chain upstream and downstream [32], and the difficulty of obtaining data beyond the supply chain [33].

To solve these problems, different frameworks/models were developed to integrate sustainability in supply chains. Zimon et al. [29] identify four categories of SSCM models/frameworks: implementation models, conceptual models, performance models, and contextual factor models. A thorough review of the different frameworks can be found in [34]. However, none of these frameworks address the possibilities of combining Big Data and Industry 4.0 to support SSCM.

2.2. Industry 4.0

Industry 4.0 is a German project that has amalgamated manufacturing with information technology, and its aim is to work with a higher level of automatization in order to accomplish a higher level of operational productivity and efficiency by connecting the physical to the virtual world [35]. This connection is achieved using technology-based production processes and equipment that communicates autonomously along the value chain [36]. Therefore, Industry 4.0 is an industrial approach based on three fundamental principles: (1) equipment and processes are connected to each other and operate with the maximum possible degree of autonomy, allowing horizontal and vertical integration across the entire value creation network; (2) digitalization of the product and service offerings and end-to-end engineering throughout the entire life cycle; (3) innovative digital business models [37].

Industry 4.0 implies complete communication among the different components of the supply chain, such as companies, factories, suppliers, logistics, resources, and customers [38]. Each of them optimizes its configuration in real time depending on the demands and status of the other members of the supply chain, which will allow for the incorporation of sustainable practices [38–42]. For example, costs and pollution, raw materials and CO₂ emissions will be reduced [43].

The information technology part of Industry 4.0 consists of cyber-physical systems (CPSs) operating in a self-organized and decentralized manner, using cloud computing and the Internet of Things (IoT) to communicate and cooperate with each other and with humans in real time [15]. This cooperation is based on Interoperability, Virtuality, Decentralization, Real-Time Capability, Modularity, and Service Orientation [43].

2.3. Big Data

Currently, there is no globally accepted definition of the term Big Data, although the 6Vs framework emerged as a structure commonly used to describe it: Volume, a very large amount of data; Velocity, the data are generated very quickly and must be processed in a very short time; Variety, a large number of structured and unstructured data types are processed; Value, the goal is to generate significant value for the organization; Veracity, reliability of the processed data; and Variability, flexibility to adapt to new data formats by collecting, storing, and processing them [44].

Big Data analytics are technologies and techniques used to analyse large-scale, complex data from various applications in order to acquire intelligence and to extract unknown, hidden, valid, and useful relationships, patterns, and information. Different methods are used to deal with such data. Some of the most important include: Text analytics, Audio analytics, Video analytics, Social Media analytics, and Predictive analytics [45,46]. Therefore, Big Data involves a complex, interconnected and multi-layered ecosystem of high-capacity networks, users, applications, and services needed to store, process, visualize, and deliver results to destination applications from different data sources [47].

Big Data analytics are being used in different aspects of SCS, such as supporting world-class sustainable manufacturing [48], designing sustainable ship routing [49] and scheduling [50] or assessing environmental efficiency [51].

3. Software Tools and Research Methodology

3.1. Software Tools

The bibliographic search was conducted using the Web of Science. Web of Science together with Scopus are the main sources of bibliographic citations used for bibliometric analyses. This is mainly because they are the only ones that combine both a rigorous selection process and a wide interdisciplinary coverage, which represents a significant strength over the others databases. On the one hand, there are other popular interdisciplinary databases such as Google Scholar, but the low data quality found in Google Scholar raises questions about its suitability for research evaluation [52]. On the other hand, there are prestigious bibliographic databases but they are focused on specialised fields, or they are regional and even country-based abstracting.

In this case, for the systematic and extensive review of Industry 4.0 and Big Data to support sustainable supply chains, relevant and high-quality papers from different scientific fields are needed, because sustainability is an interdisciplinary field addressed from different research domains, such as business management, environmental and social sciences, engineering, etc. Therefore, Web of Science together with Scopus are the best databases for this study.

Even so, both have their limitations and biases that give rise to the existence of underrepresentation of articles by fields, language, or countries. The most comprehensive approach would be to combine different search engines so that they can each cover the biases of the others [52]. In the study presented in this paper only the Web of Science was used because it offers several analysis tools that are necessary and that are not compatible with other systems. The reason for not using Scopus is that its search tool works by study areas and does not offer the versatility to carry out the search by content, title or author keywords, which is a crucial element for this study.

The desktop version of EndNote was used for the management of the bibliographic libraries extracted from the search in the Web of Science. This software was selected because it is a tool that is very easy to use, very powerful, compatible with many bibliographic search engines, compatible with the MS Office tools and can be downloaded as a trial version.

A tool integrated within the Web of Science was used for the bibliographic analysis. This tool performs an analysis of the list of articles, grouping them into different categories, such as country, institution in which they were developed, the journal that published them or their authors. It also allows a report to be created that provides an analysis of the evolution of the number of citations over time and obtains some reference indicators such as the h-index.

The list of articles was extracted from the Web of Science in plain text format and was then processed with the BibExcell tool to generate a file with the necessary format and structure for the network analysis performed in Gephi. Gephi is an open source tool widely used for network and graph analysis. It offers the possibility of performing an advanced exploratory analysis of entity–relationship data. It also allows a significant number of distribution, organization, and clustering algorithms to be applied in order to determine the interactions that occur in the data. These techniques allow, among other things: identification of nodes, or more influential entities; identification of groups related by affinity; calculation of relevance indicators such as PageRank; or visualization of information in order to facilitate the identification of elements of interest.

3.2. Research Methodology

The model for the methodology followed to perform the bibliographic analysis was taken from the article by Mishra et al. [19], which in turn was based on the five-step method proposed by Rowley and Slack [17].

First, the keywords that defined the search for the articles were selected. Three keyword combinations were established: (1) Big Data, Industry 4.0, sustainability, supply chain; (2) Big Data, sustainability, supply chain; and (3) Industry 4.0, sustainability, supply chain. The search was conducted in English as there are a greater number of bibliographic sources in that language [35]. Web of Science

has the search criterion 'Topic', which searches for the keyword combinations in the title, abstract, and keywords of the papers. This is the search criterion that was used in the study.

The search method used in this study is based on the snowball system that is widely described in Wohlin's article [18]. As a result, an initial list of articles was obtained from the combination of the aforementioned keywords. The articles in the initial list were then filtered by analysing each of them individually. First, we checked whether at least two of the search keywords were contained among the keywords of the article in question. If that was the case, references to the third keyword were sought in the abstract. If they were not included, the abstract was analysed in detail to determine whether the article dealt with the topic in question, even if the keyword was not explicitly mentioned. If the keyword was not found, the article was discarded. If it was found, the article was selected, and then its cited papers were analysed to determine whether they could also be selected. If so, the process was continued in a similar manner until no further adequate articles were found in the references of a selected article. When this happened, it returned to a higher level to continue the search.

Applying this process resulted in a list of 87 articles. Once the list was defined, the analysis tools available on the Web of Science were used. The following tasks were performed with these tools:

- Analysis of the evolution over time of the number of articles published included in the list,
- Analysis of the evolution of the number of citations generated by the articles,
- Analysis of the number of articles published by author,
- Analysis of the number of articles published by country,
- Analysis of the number of articles published by institution,
- Analysis of the content of the 10 most cited articles on the list,
- Analysis of the h-index indicator,
- Analysis of the number of articles published per journal,
- Analysis of the indicators of relevance, impact, and prestige of the 10 journals with the most published articles on the list. The indicators analysed were the following: CiteScore, Impact Factor, Normalized Source Factor, and Scimago Journal Rank.

The plain file containing the list of papers was treated with BibExcell to generate a file that was compatible with Gephi software. This file contained only the relationships among all the articles and a label that identified them. The relationships reflected the citations that papers on the list made from other papers on the list. The analysis of networks and relationships carried out with Gephi was aimed at: (1) analysing the relationship of articles based on citations made among them; (2) identifying the relevant articles and those that were marginal; (3) detecting clusters or sub-communities on the list based on the relationships generated through citations; and (4) calculating the PageRank indicator.

PageRank was the first algorithm used by Google to establish a ranking among web pages, proposed by Brin and Page [53]. The algorithm assigns a value to each web page based on a network analysis that measures the interactions among the pages. This algorithm can be applied to any set of elements that are related to each other through citations or references. This is why, shortly after its appearance, it was applied in bibliographic studies and in determining the relevance and prestige of publications [54]. The interesting thing about this algorithm is that it not only takes into account the times an article is cited but also the degree of importance of the articles that cite it.

Finally, a detailed analysis of the content of all the articles was carried out in order to classify them in different categories. A system inspired by the comparative method proposed by Collier [20] was used to establish the categories. According to this system, a preliminary classification based on a previous content analysis was proposed. To establish these categories, the common points shared by the articles were identified. The fundamental objective of the article, as well as its contributions and the advances it offered to the state of the art were the main factors considered. This classification was taken as a starting hypothesis. After that, the adequacy of the categories to classify all the articles was checked paper by paper. When an article was found that did not fit in any category, the classification

was rethought with a view to integrating the dissonant element. Several reviews were performed until all the items on the list were properly distributed in the proposed classification.

4. Bibliometric Analysis

4.1. Initial Results

First, it can be said that the number of articles on the final list was relatively low when compared with specific searches for only one of the keywords. For example, on the Web of Science 54,371 articles appeared for the keyword 'supply chain' and 56,546 for 'Big Data'. However, combining the two keywords reduced the number to 679 papers. When the search was performed with the three keywords in one of the combinations selected for the study (Big Data, supply chain, and sustainability), 93 results appeared, representing 13.7% of the 679 selected in the previous search that did not include the word 'sustainability'. From this, it can be deduced that an important number of the studies related to Big Data as applied to the supply chain took into account the dimension of sustainability, which is a relevant fact. These searches were carried out as of June 2019. Finally, it should be clarified that, although the search that was performed with the words 'Big Data', 'supply chain', and 'sustainability' offered 93 results, several were left out of the final selection because of the filtering process described in the methodology section.

The analysis tools of Web of Science were used to carry out a global analysis of the list of articles obtained. Attention should be paid to the clear upward trend that can be observed in relation to the number of articles generated on this subject in the years covered by the study (Figure 1). In the case of 2019, since it was not complete, a projection was carried out using linear regression. The result was 36 articles, with a 95% confidence interval.

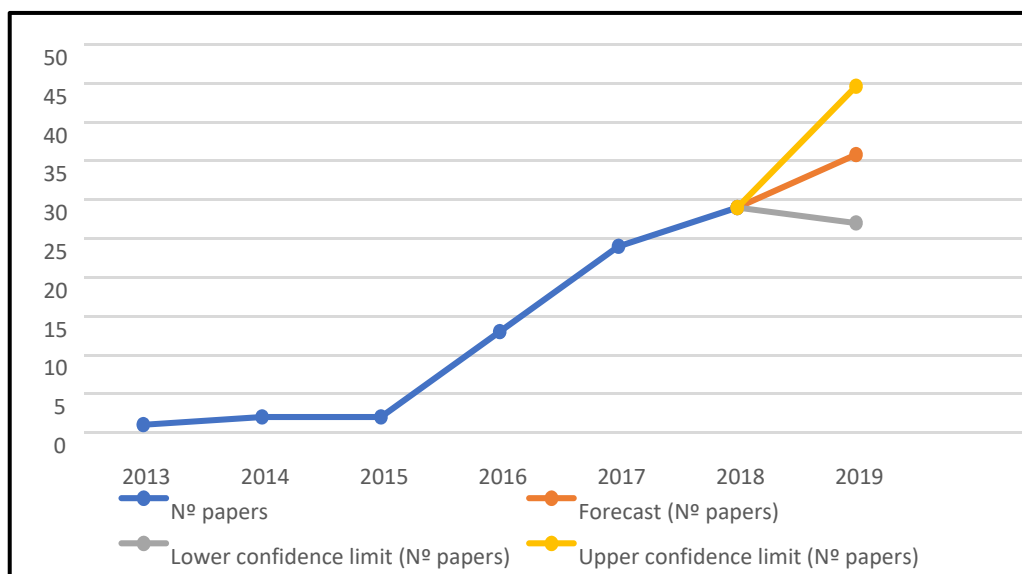


Figure 1. Trend in the generation of articles. Source: Web of Science.

This trend can also be observed in relation to the number of accumulated citations (Figure 2). In the case of 2019, a projection was carried out using linear regression. The result was 721 citations, with a 95% confidence interval.

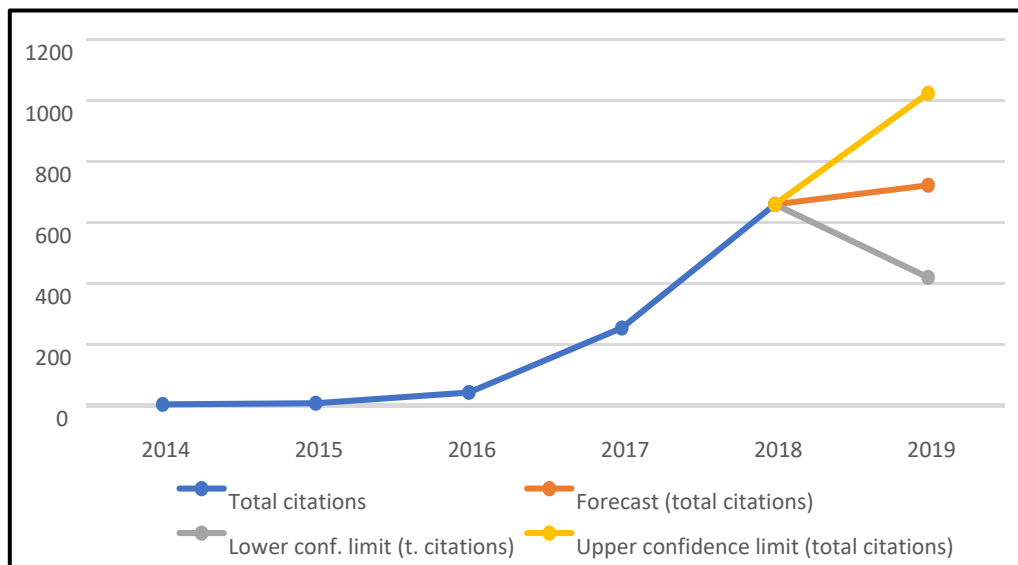


Figure 2. Evolution in the total citations. Source: Web of Science.

4.2. Author Influence

There are several authors who stood out above the rest in relation to the number of publications written on the subject. In particular, A. Gunasekaran and T. Papadopoulos each published five articles during the analysis period. Table 1 shows the 10 most relevant authors.

Table 1. Ten authors with the most articles. Source: Web of Science.

Author	Total Papers
Gunasekaran A.	5
Papadopoulos T.	5
Childe S.J.	4
Dubey R.	4
Hazen B.T.	4
Akhtar P.	3
Khan Z.	3
Liu Y.	3
Rao-Nicholson R.	3
Boone C.A.	2

4.3. Affiliation Statistics

Another issue that stood out from the analysis was the distribution by countries. The United States and United Kingdom together accounted for almost 50% of the items generated. Behind them China and India represented the Asian countries. Together, these four countries covered the major part of the studies. From a continental point of view, the Europeans researched the most in this field, followed by the Asians (Table 2).

Table 2. Ten countries with the most articles. Source: Web of Science.

Country	Total Papers
United States of America	22
United Kingdom	21
China	18
India	10
Taiwan	7
France	6
Finland	5
Germany	4
Brazil	3
Italy	3

4.4. Analysis by Institution

The three institutions with the largest number of articles published were the English universities of Kent, Plymouth, and Hull. Table 3 shows the 10 institutions with the largest number of published articles.

Table 3. Ten institutions with the most articles. Source: Web of Science.

Institution	Total Papers
University of Kent	6
University of Plymouth	5
University of Hull	4
University of Tennessee Knoxville	4
University of Tennessee System	4
California State University	3
Dalian University of Technology	3
Languedoc-Roussillon Universities	3
Montpellier Business School	3
Symbiosis International University	3

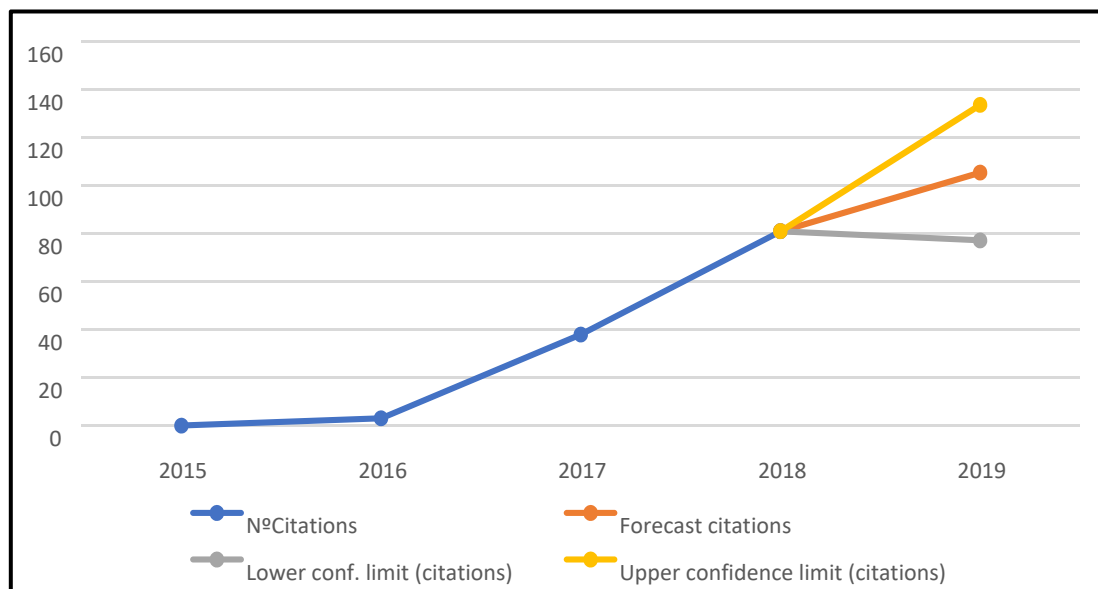
4.5. Citation Analysis

With regard to the 10 most cited articles (Table 4), only two were bibliographic analyses. Most of them were models or theoretical developments with a practical application, and in one way or another served to facilitate the incorporation of new data analysis technologies to specific aspects of supply chain management.

In first position, with a total of 158 citations, was the article by Wang, Gunasekaran, Ngai and Papadopoulos [9]. This article carries out a review of the literature that deals with the application of data analytics technologies, including Big Data, to the management of supply chains considering their sustainability. They also develop the use of data management techniques in that context. Figure 3 shows the evolution of the citations since the publication of the article. It can be observed how the trend ascends quite sharply. Given the relatively recent publication of the paper, the number of citations was very high compared to most of the papers on the list. In the case of 2019, a projection was carried out using linear regression. The result was 105 citations, with a 95% confidence interval. The total number of citations mentioned above (158) does not take into account the projection, only the actual citations at the time of the study (June 2019).

Table 4. Ten articles with the most citations. Source: Web of Science.

Paper	Total Citations
Wang, Gunasekaran, Ngai and Papadopoulos [9]	158
Shrouf, Ordieres and Miragliotta [55]	123
Chae [56]	92
Dubey, Gunasekaran, Childe, Wamba and Papadopoulos [57]	77
Zhang, Ren, Liu and Si [58]	76
Papadopoulos, Gunasekaran, Dubey, Altay, Childe and Fosso-Wamba [59]	57
Zhao, Liu, Zhang and Huang [60]	53
Wu, Liao, Tseng, Lim, Hu and Tan [61]	52
Fawcett and Waller [62]	50
Ur, Chang, Batool and Wah [63]	44

**Figure 3.** Evolution of article citations of Wang et al. (2016). Source: Web of Science.

In second position was the article by Shrouf, Ordieres and Miragliotta [55]. This article had 123 citations and establishes a reference architecture for the development of Industry 4.0 projects in the context of the so-called Smart factories, with a focus on sustainability. It also deals in depth with energy management. Figure 4 shows the evolution of the citations of this article over time. In the case of 2019 a projection was carried out using linear regression. The result was 69 citations, with a 95% confidence interval. The total number of citations mentioned above (123) does not take into account the projection, only the actual citations at the time of the study (June 2019). Here it can be observed that the lower and upper confidence limits are very close to the prediction; this is due to the fact that the dating trend followed a path close to a straight line and was therefore easier to predict. In this case, unlike the previous one, the number of citations was initially much higher (16 versus 3 for the previous one) but from there on, the rise was much gentler for the second case.

In third position was the article by Chae [56], which had 92 citations. This article applies Big Data techniques such as sentiment analysis to generate knowledge related to issues like corporate social responsibility or human rights in the field of supply chain and logistics. From there, the potential of social networks for practices related to supply chain management is evaluated. The evolution of the citations for this case is shown in Figure 5. For 2019, a projection was carried out using linear

regression. The result was 48 citations, with a 95% confidence interval. The total number of citations mentioned above (92) does not take into account the projection, only the actual citations at the time of the study (June 2019).

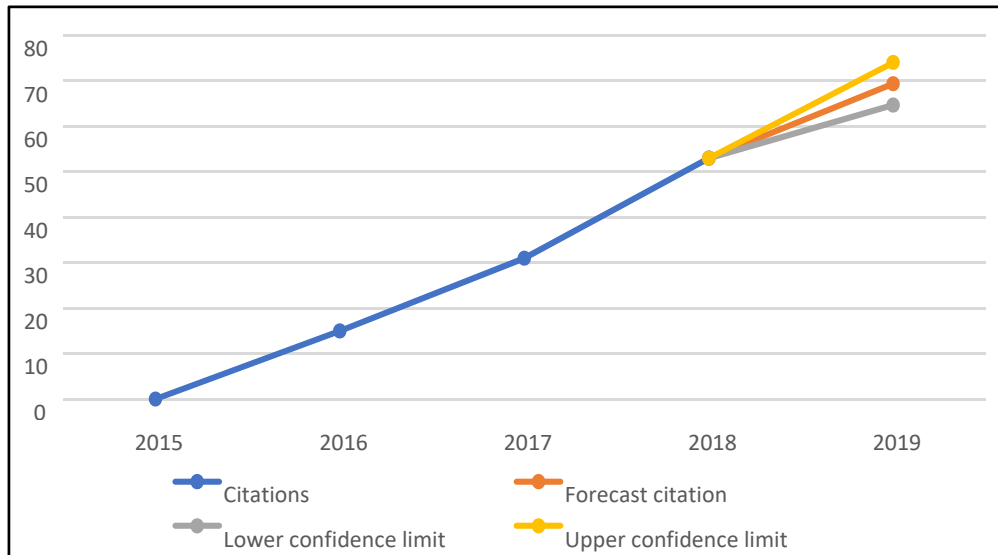


Figure 4. Evolution of article citations of Shrouf et al. (2014). Source: Web of Science.

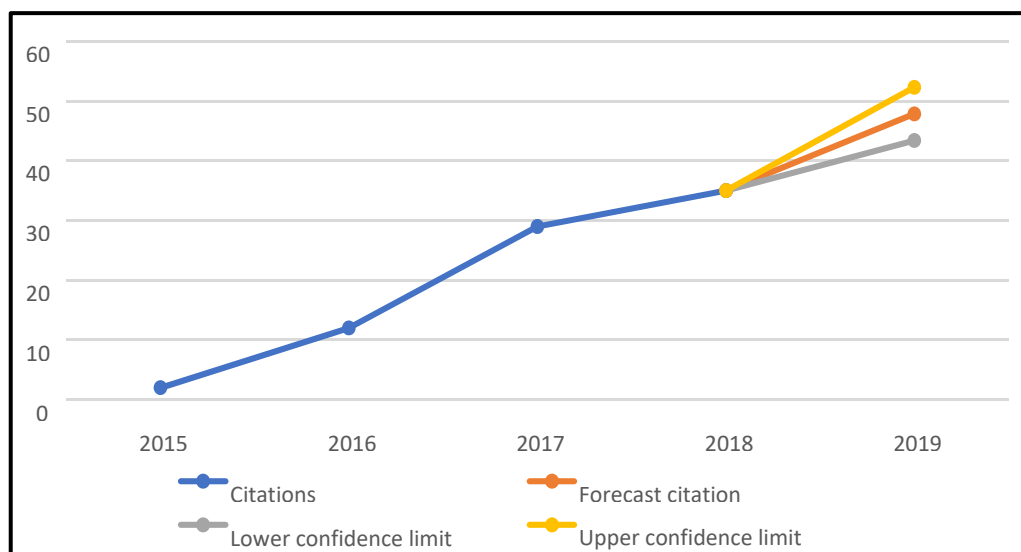


Figure 5. Evolution of article citations of Chae (2015). Source: Web of Science.

In fourth position was the article by Dubey, Gunasekaran, Childe, Wamba and Papadopoulos [57], which analyses the role of Big Data analytics in supporting global scale manufacturing systems. This analysis includes structured interviews with senior managers and concludes with a proposal for a conceptual model that is tested with the data obtained throughout the study.

In fifth position was the article by Zhang, Ren, Liu and Si [58], which proposes an architecture for Big Data analysis applied to the product life cycle. The proposed architecture is tested in a case study.

In sixth place was the article by Papadopoulos, Gunasekaran, Dubey, Altay, Childe and Fosso-Wamba [59], which proposes and tests a framework to explain resilience in supply chain networks using Big Data analytics on data extracted from social networks, among other sources. The article draws its conclusions by analysing the context of the earthquake that occurred in Nepal in 2015.

In seventh place was the article by Zhao, Liu, Zhang and Huang [60], in which a multi-objective optimization model is developed for the management of sustainable supply chains and for which they apply Big Data.

In eighth place was the article by Wu, Liao, Tseng, Lim, Hu and Tan [61], which proposes a method to use Big Data analytics to determine causal relationships of risks and uncertainties in supply chains, integrating sustainable indicators, among other things.

In ninth place was the article by Fawcett and Waller [62], which analyses five elements with potential to revolutionize the design of supply chains, one of which is Big Data. Four barriers that prevent high levels of co-creation of value among members of the supply chain are also identified, one of which is the poor understanding of corporate social responsibility initiatives.

In tenth place was the publication by Ur, Chang, Batool and Wah [63], which presents a framework for the best management of Big Data in the context of sustainable companies, with the ultimate goal of generating value for the company.

4.6. H-index

Web of Science provides an impact meter, the h-index. The h-index is an indicator proposed by [64] which, according to the author, is better than others that are commonly used to measure the impact of a researcher, such as Number of articles published, Total number of citations, Citations per article, Number of significant articles (establishing as significant those that are based on a minimum number of citations), or Number of citations of the n most cited articles. In fact, all these measures have disadvantages. For example, the number of citations per publication rewards low productivity (number of articles published), while the number of articles published, in contrast, rewards productivity without taking into account their quality. This issue is problematic because, depending on how scientific publications are evaluated, research is conducted in one way or another. The use of simple quantitative meters (which do not meet the quality of the article) is negatively affecting the generation of articles with long-term impact. This topic is dealt with in depth in articles such as Smith et al. [65].

In response to this problem, indicators such as the h-index emerged. This index corresponds to the number n of articles that were cited at least n times. The h indicator of the list of 87 articles that resulted from the literature search was 19, so that there were 19 articles that were cited at least 19 times. The indicator was calculated by ordering the items on the list by the number of times they were cited (from highest to lowest) and the list was scrolled until the number of citations was less than or equal to the position of the item on the list. This comprised the h-index.

4.7. Sources Analysis

Table 5 shows the 10 journals with the highest number of articles on the list. The most commonly used impact and relevance indicators are also included. It can be noted that in general terms they are sources that address the application of sustainable policies in the industrial or logistics sector. An example is the first, *Journal of Cleaner Production*, published by Elsevier. It addresses in depth many aspects related to sustainability, such as governance, the environment, corporate social responsibility and how they can be introduced into production systems.

The indicators reflected in Table 5 express the degree of impact, relevance and importance of the journal. There was a great deal of debate about which indicators were best suited to measure the impact of a journal. The relevance indicators shown in Table 4 are the most used and accessible. They are briefly explained below:

- i. CiteScore: Measures the average number of citations received per document published in the journal. Values are calculated by counting citations over a year for documents published in the three years prior to the calculation and dividing by the number of documents published in those three years. As a comparative reference to the results shown in Table 4, the best score for the year 2018 was 160.19 and the average value was 1.6337.

- ii. **Impact Factor:** This is another widely used impact meter. The difference with respect to the previous one is that, instead of taking the publications of the three previous years, it does it with a time range of two years. The best score for 2018 was 244.585.
- iii. **Source Normalized Impact per Paper (SNIP):** This index measures the impact of citations in a given context. It is based on total citations per field of study. The impact of a citation has a greater value in fields where citations are less likely to occur. The best score for 2018 was 100.014 and the average value was 0.8566.
- iv. **SCImago Journal Rank (SJR):** This measure takes into consideration the prestige of the journal in which the article is published. It uses an algorithm similar to Google to establish rankings between websites. It also takes into account the citations of the article. The best score for 2018 was 72.576 and the average value 0.7244.

Table 5. Ten journals with the most published articles and their impact indicators. Source: Web of Science and official websites of the publications. SNIP: Source Normalized Impact per Paper. SJR: SCImago Journal Rank

Journal	Total of Papers	CiteScore	Impact Factor	SNIP	SJR
<i>Journal of Cleaner Production</i>	13	7.32	5.651	2.308	1.62
<i>Sustainability Journal</i>	10	3.01	2.592	1.169	0.55
<i>Computers Industrial Engineering</i>	4	4.68	3.518	1.755	1.334
<i>International Journal of Logistics Management</i>	4	3.28	2.226	1.134	0.871
<i>International Journal of Production Economics</i>	4	7.13	4.998	2.486	2.475
<i>Production Planning Control</i>	3	4.38	3.340	1.514	1.43
<i>International Journal of Production Research</i>	3	4.34	3.199	1.720	1.585
<i>Process Safety and Environmental Protection</i>	3	4.60	4.384	1.626	1.075
<i>Journal of Manufacturing Systems</i>	2	5.45	3.642	2.234	1.592
<i>International Journal of Physical Distribution Logistics Management</i>	2	6.60	5.212	2.109	2.41

4.8. Data Clustering Using Content Analysis

A content analysis of all the articles on the study list was carried out to define categories that classify articles based on common elements, in order to bring some order to the research effort that was being made, and to identify future research suggestions in Big Data, SSCM, and Industry 4.0. The categories obtained are the following:

1. *Applied research:* This category includes all the articles whose objective is to develop a framework, model, or system that can be used in some practical context to solve a problem that has been detected. The proposal is validated through its application to a case study.
2. *Diagnosis:* This category encompasses articles that perform a purely theoretical analysis of the status or evolution of a given theme or area of study. The most influential elements are identified and possible future evolutions, patterns, principles, etc., are established.
3. *Bibliographic study.* This category includes articles that perform a review of the published literature that addresses a subject or area in question (usually bounded by keywords). Among other aspects, the number of published articles and their impact and trends over time are analysed. In addition, it also identifies the elements in which more interest is shown and knowledge gaps. The ultimate goal is to give a complete diagnosis of the state of the art in order to influence the trends detected.
4. *Impact analysis.* This category includes articles that analyse the impact that an element has on a real phenomenon. It is a practical application focused on a specific case. The impact referred to is evaluated and contrasted with data and conclusions are drawn based on the results.
5. *Theoretical postulate:* This category includes articles that revolve around the argumentation and foundation of a theoretical proposal that does not constitute a framework for practical application, but instead moves in the field of principles, foundations, and relevant elements linked to an area of study or a phenomenon. No concrete proposal is made that has any practical application.

6. *Specific solutions.* This category comprises articles that present a practical solution to a very specific problem. Apart from explaining the main points of the proposed solution, its functionality is contrasted with real case studies. Within this framework, different programming models, algorithms, or indicators can be found.

Table 6 presents the distribution of the 87 papers on the list in each category, and a compilation of the future research suggestions in Big Data, SSCM, and Industry 4.0 made by a content analysis of the papers in every category.

Table 6. Total papers per category.

Categories	Papers	Top 5 Papers	Future Research Suggestions
Applied research	26	Shrouf, et al. [55]; Zhan et al. [58]; Dubey [57]; Zhao et al. [60]; Ur et al. [63]	To embrace Big Data Analytics to redefine the future focus of the advanced manufacturing technology in SSCM, considering new innovations as, for instance, developing new materials such as biodegradable materials
Diagnosis	34	Fawcett [62]; Wu et al. [61]; Sanders et al. [66]; Song et al. [67]; Kusiak [68]	Application-oriented research; Methodologies and tools to support managers to improve the efficiency and effectiveness of the decision-making process; To analyse the relationship between firms' capabilities and Industry 4.0 and SSCM
Bibliographic study	7	Wang et al. [9]; Cerchione and Esposito [69]; Tiwari et al. [70]; Kamble et al. [15]; Kuo and Smith [71]	Collaboration between enterprises and academic researchers to develop new eco-innovative technologies to meet their specific needs; To analyse the factors affecting the adoption of knowledge management practices
Impact analysis	7	Papadopoulos et al. [59]; Zhong et al. [72]; Rodriguez and Da Cunha [73]; Mani et al. [16]; Hopkins and Hawking [74]	Guidance regarding how organizations might identify future versus present needs; To increase the number of case studies
Theoretical postulate	5	Lopes de Sousa et al. [12]; Richey et al. [75]; Jensen and Remmen [76]; Liu and Yen [77]; Keivanpour and Kadi [78]	To understand the connections, integration, and mutual benefit between environmentally sustainable manufacturing and Industry 4.0; The Key Success Factors must be developed by researchers, policy makers, and industrialists together; To carry out qualitative research and to convert the research postulates into hypotheses that should be tested through quantitative methods
Specific solutions	8	Chae [56]; Hazen et al. [79]; Badiyadeh et al. [80]; Shen and Lam [81]; Banyai et al. [82]	To understand social media and social media data in SSCM; To analyse behavioural and marketing-related issues; To analyse consumer perceptions of remanufactured products; To assess SSCM in the presence of fuzzy and stochastic data

4.9. Network Analysis: Gephi

Finally, a network analysis was made using Gephi. It showed the relationships among the papers on the list, such as the paper most cited by other papers on the list, and the calculation of the PageRank indicator. The results are showed in the Appendix A.1.

5. Discussion

5.1. Contributions to Theory

The current study contributes to the literature on Big Data, Industry 4.0, and SSCM and extends current reviews [9,19,68,71] in that: (1) it identifies which are the top contributing authors, countries, and institutions in the field of Big Data and Industry 4.0 applied to SSCM, using both statistical analysis and techniques of bibliometric and network analysis to obtain and compare the most influential works (based on citations, co-citations, and PageRank) (answer to RQ1); (2) through a content analysis, it identifies and proposes six research categories (applied research, diagnosis, bibliographic study, impact analysis, theoretical postulate, and specific solutions), that focus on particular areas, from bibliographic and conceptual studies to methods, tools, and case studies of Industry 4.0 and Big Data applied to SSCM (answer to RQ2); and (3) it identifies future research necessities in the field of Big Data and Industry 4.0 applied to SSCM (answer to RQ3).

Therefore, this work covers an important research gap: the scarcity of systematic and extensive reviews of the recent research on Industry 4.0 and Big Data applied to SSCM, which could limit its impact [6,16]. As a result of the study, some conclusions can be drawn regarding the literature focused on sustainability, supply chains, Big Data, and Industry 4.0: First, it was seen how the number of publications is still limited. However, a significant percentage of them (13.7% of publications in

the field of Big Data and supply chains) cover the issue of sustainability. It was also noted that the trend is clearly rising, both in the number of publications per year and in the number of citations per year. From this it can be deduced that there is a growing interest in this area, which is in line with the work of [9]. Second, in terms of productivity, A. Gunasekaran and T. Papadopoulos were identified as the authors with the most publications in the field of study, and United Kingdom and USA were the countries that contributed the greatest number of publications. Third, the h-index of the list of 87 papers is 19, which is somewhat low compared with the median calculated by Malesios and Psarakis [83] for the fields of Computer Science (median of 20) and Social Sciences (median of 40), which are the ones that are best related to the field considered in this study. Fourth, the journals where the 87 articles on the list were published have a higher than average impact in their area, but their impact is much lower than the values obtained by the most relevant journals. Fifth, from the network analysis, the most important conclusion that can be drawn is that the degrees of relationship among articles through citations does not provide a good basis for identifying research categories or clusters. Sixth, a comparative analysis of the content made it possible to put forward a classification by categories that include all the papers on the list. This classification in six research categories facilitates the future work of researchers interested in this field because it identifies common shared patterns and elements in every category, and shows those aspects that were addressed to a lesser extent (as is the case of those categories with fewer publications).

Finally, other important contributions of this work are that it can be the basis for two research challenges in the field. On the one hand, this bibliographic analysis can be extended to other technologies of great potential, such as machine learning, artificial intelligence, or cloud computing. This would further complete the research and would facilitate the incorporation of other promising technologies in SSCM, since many of these technologies are complementary [84]. On the other hand, as none of the existing SSCM frameworks/models address the possibilities of combining Big Data and Industry 4.0 to support SSCM, a new framework/model to cover this research gap can be developed based in the results of this systematic literature review. It was already proved that conceptual frameworks/models can be derived from systematic literature reviews [29]. For example, in the SSCM field, Gosling et al. [14] developed a framework that integrates supply chain leadership and supply chain learning perspectives; Rebs et al. [26] proposed a framework to apply system dynamics modelling for sustainable supply chain management; Zimon et al. [29] analysed how to cover the gap between corporate sustainability strategy and SSCM implementation, proposing three possible models (Reactive, Cooperative, and Dynamic Models); Svensson [85] developed a conceptual framework with an empirical example to describe and illustrate aspects of the first-, second- and n -order supply chains; and Zimon et al. [28] analysed the alignment of the supply chain with the United Nations Sustainable Development Goals.

5.2. Contributions to Managerial Practice

This work offers different opportunities to practitioners. This study can offer supply chain managers and consulting firms different schools of thought that will enable them to take advantage of the benefits from applying Big Data and Industry 4.0 to SSCM. Furthermore, through the classification of the literature in six categories, practitioners can: (1) assess the current state of the art in Big Data and Industry 4.0 applied to SSCM, in terms of conceptualisation, methods, tools, impact, specific solutions, and case studies; (2) identify their future requirements in the six categories to make appropriate decisions on whether to invest and improve current tools/methods; and (3) analyse the implications of Big Data and Industry 4.0. for the achievement of their SSCM strategy.

6. Conclusions

In this paper, a bibliographic analysis of the literature on Big Data, Industry 4.0, and supply chain management published since 2009 was carried out. A total of 87 papers were analysed in order to identify the evolution over time of the number of articles published that are included in the list; the evolution of the number of citations generated by these articles; the number of articles published

by author; the number of articles published by country; the number of articles published by institution; the content of the 10 most cited articles on the list; their h-index indicators; the number of articles published per journal; the indicators of relevance, impact, and prestige of the 10 journals with the most published articles on the list; the established and emerging research categories on the topic; and the relationships among the 87 papers.

The bibliographic analysis proved the initial hypothesis that an analysis of current research can facilitate the advancement of future research in this area. The main conclusion is that the area of study requires more research and a higher number of annual publications. It is also necessary to improve the relevance of the research carried out, something that can be achieved by accessing journals of greater impact. Finally, the research conducted in four of the six identified categories should also be improved, since the majority of papers were published in the Applied Research and Diagnosis categories.

Finally, it is important to highlight the limitations of the study. This study was limited mainly by (1) the biases introduced by studying a single bibliographic database, the Web of Science, which has shortcomings in terms of publications in the field of humanities or certain social sciences. There was also a language bias, since this database includes mostly articles that were written in English, and the search was conducted only in English. Other important databases, such as Scopus Database, could be used to improve and compare the results; (2) choosing a series of specific keywords introduced another bias by default. For example, other keywords, such as sustainable supply chain management could have been used and might have yielded different results; (3) the bibliometric and network analysis for reviewing the literature based on [19] was used. Other methods might be used for such an analysis; and finally, (4) the literature was classified in six research clusters. Other methods may result in other classifications.

Author Contributions: R.C.: Conceptualization, supervisión, future research identification, writing-original draft preparation, review and editing. N.J.S.-d.: methodology, bibliometric analysis. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: Authors declare no conflict of interest.

Appendix A

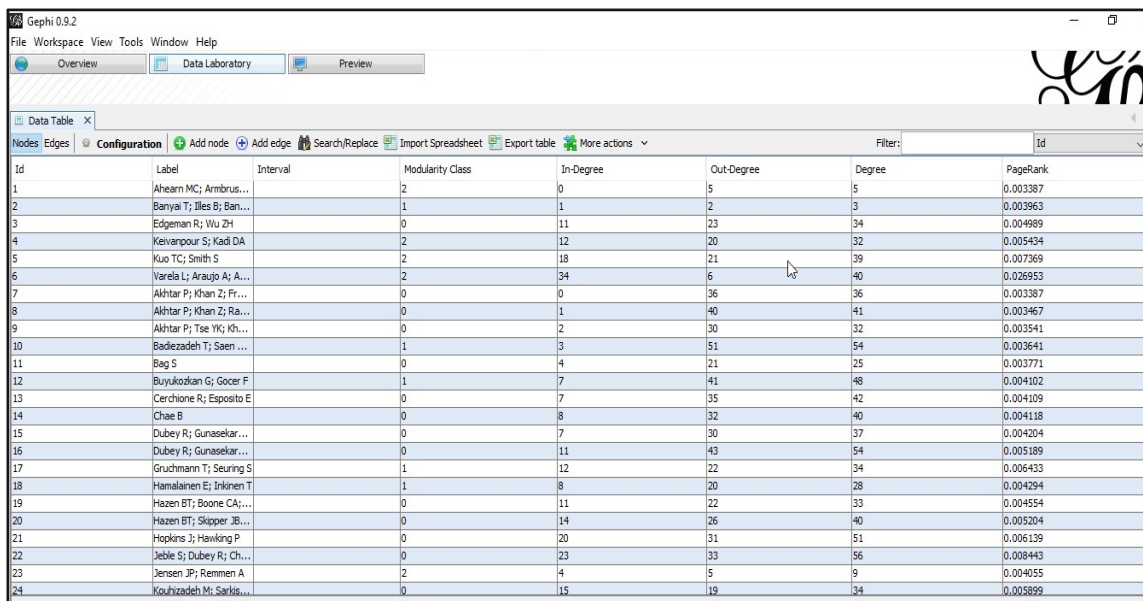
Appendix A.1 Network Analysis: Gephi

Appendix A.1.1 Initial Analysis

On entering the list of articles in the Gephi program, it created 85 nodes (each node corresponds to an article). The reason why there are two excluded items is because they do not cite any other papers from the list, nor are they cited by another item on the list. The program found 1192 relationships through citations among the items on the list. The list entered in the program is displayed in table format. Figure A1 shows a screen capture of this table generated by the Gephi program.

The Force Atlas algorithm was applied. This places the most cited articles in the middle and those less cited on the periphery. A colour gradient was added to darken the circles, depending on the number of times the article was cited. Following these two applications, the relationship diagram is as follows (Figure A2).

If the mouse cursor is positioned on top of one of the nodes, the number of papers that are related to it is displayed. Figure A3 shows an example corresponding to the article that had the highest number of interactions and was cited the most times: Tiwari et al. [52]. Oddly, this article was not the most cited in general but it was the most cited by the articles on the list, which established an interesting distinction. Within the specific field of research, this article was the most relevant and therefore should have priority over others that, despite having more citations, were from articles belonging to other fields of study.



Id	Label	Interval	Modularity Class	In-Degree	Out-Degree	Degree	PageRank
1	Ahearn MC; Armbros...		2	0	5	5	0.003387
2	Banyai T; Iles B; Ban...		1	1	2	3	0.003963
3	Edgeman R; Wu ZH		0	11	23	34	0.004989
4	Keivanpour S; Kadi DA		2	12	20	32	0.005434
5	Kuo TC; Smith S		2	18	21	39	0.007369
6	Varela L; Araujo A; A...		2	34	6	40	0.026953
7	Akhtar P; Khan Z; Fr...		0	0	36	36	0.003387
8	Akhtar P; Khan Z; Ra...		0	1	40	41	0.003467
9	Akhtar P; Tse YK; Kh...		0	2	30	32	0.003541
10	Badiezadeh T; Saen ...		1	3	51	54	0.003641
11	Bag S		0	4	21	25	0.003771
12	Buyukozkan G; Gocer F		1	7	41	48	0.004102
13	Cerchione R; Esposito E		0	7	35	42	0.004109
14	Chae B		0	8	32	40	0.004118
15	Dubey R; Gunasekar...		0	7	30	37	0.004204
16	Dubey R; Gunasekar...		0	11	43	54	0.005189
17	Gruchmann T; Seuring S		1	12	22	34	0.005433
18	Hamalainen E; Inkinen T		1	8	20	28	0.004294
19	Hazen BT; Boone CA;...		0	11	22	33	0.004554
20	Hazen BT; Skipper JB...		0	14	26	40	0.005204
21	Hopkins J; Hawking P		0	20	31	51	0.006139
22	Jebble S; Dubey R; Ch...		0	23	33	56	0.008443
23	Jensen JP; Remmen A		2	4	5	9	0.004055
24	Kouhizadeh M; Sarbis...		0	15	19	34	0.005892

Figure A1. Data laboratory section of the Gephi tool.

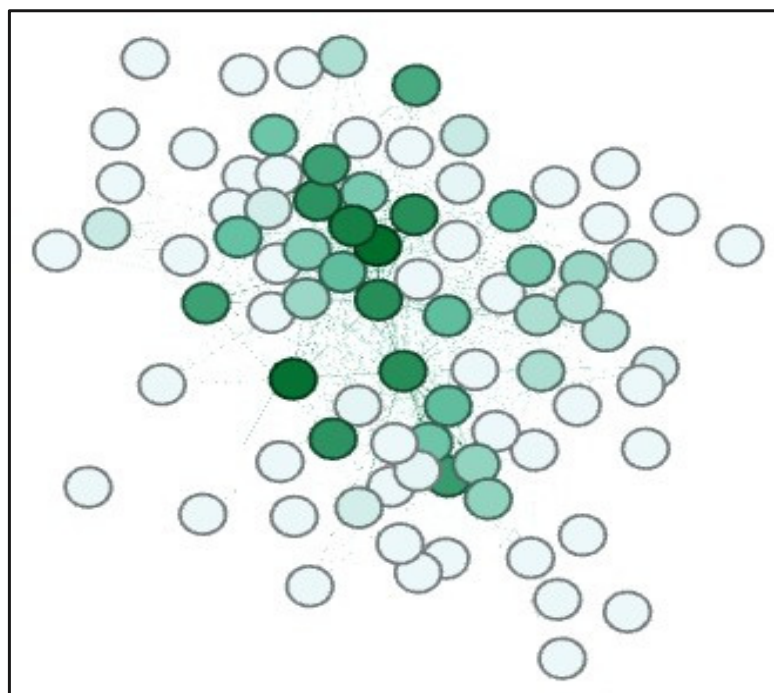


Figure A2. Diagram of the relationships among the papers on the list.

The Middle Grade method, which is one of the statistical methods available in the tool, was used to calculate the number of times an article was cited by others on the list and the number of times the article cited others. In this regard, the article Tiwari et al. [70] was cited 44 times and it cited another one eight times. In comparison, as a result of applying the same method to one of the periphery nodes (items that were rarely cited by elements on the list), Alfian et al. [86], the number of times the article was cited was zero and the number of articles on the list that were cited was four.

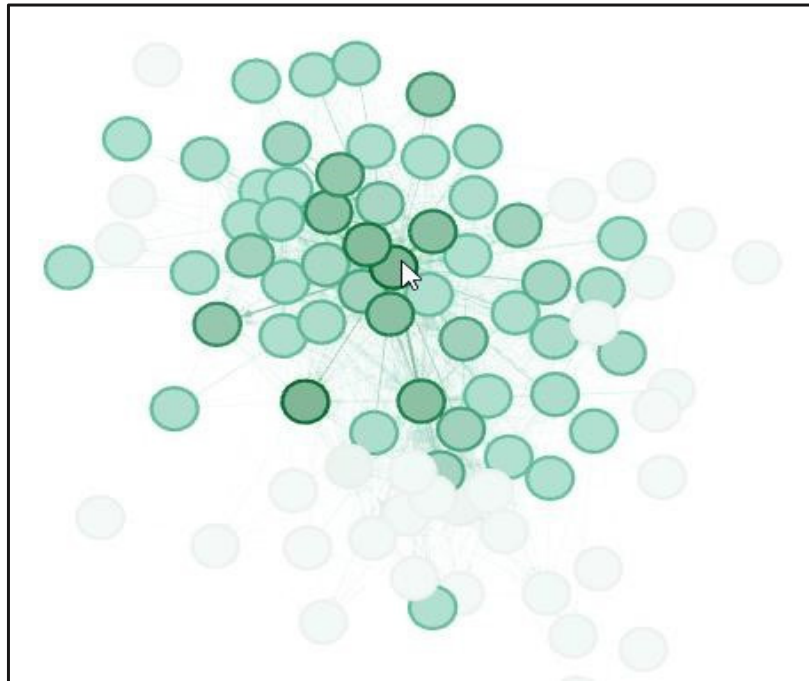


Figure A3. Paper most cited by other papers on the list.

Appendix A.1.2 PageRank

Another element of great interest in Gephi is that it allows the calculation of the PageRank indicator. It assesses the number of times that a paper is cited by other highly cited papers. Figure A4 shows the calculation of PageRank for paper A. All the nodes that cite node A, which are in turn divided by the factor $L(X)$ that corresponds to the number of citations that node X makes in relation to other nodes, are shown in brackets. The parameter d is known as the dumping factor, and serves to deal with those nodes that do not cite any other node. The formula presented here is a variant of the original that makes all values remain between 0 and 1. The value chosen to be assigned to parameter d in this study is the one that comes by default in the Gephi tool (0.85) and that is recommended in the original publication [53]. The algorithm was applied and then a colour was assigned based on the result. Nodes that are coloured are the ones with the highest rating. The colour caption is shown in Figure A5.

$$PR(A) = \frac{1-d}{N} + d \left(\frac{PR(B)}{L(B)} + \frac{PR(C)}{L(C)} + \frac{PR(D)}{L(D)} + \dots \right)$$

Figure A4. PageRank equation applied.









	0.0033868398438594125
	0.0034475321614728777
	0.003465685529370744
	0.0034668276463269312
	0.003540515190528239
	0.0035562257784965104
	0.0036169180961099757
	0.0036408514934083386

Figure A5. Colour caption.

Figure A6 shows the distribution obtained after applying the Noverlap distribution algorithm to prevent the nodes from overlapping, and thus they can be seen more clearly. Table A1 shows the top 10 papers with the highest PageRank scores.

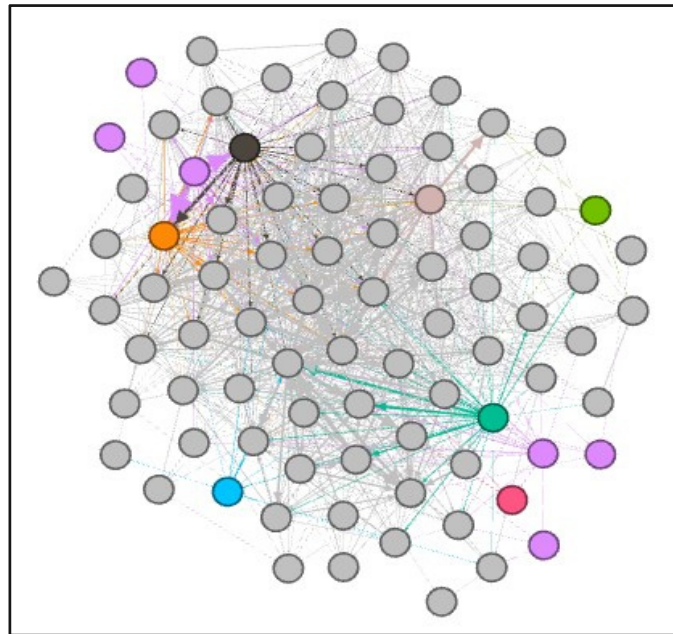


Figure A6. Distribution after applying Noverlap.

Table A1. Papers with the highest PageRank scores.

Paper	PageRank
Ahearn et al. [87]	0.003387
Akhtar et al. [88]	0.003387
Alfian et al. [86]	0.003387
Ardito et al. [89]	0.003387
Coble et al. [90]	0.003387
Ji and Sun [91]	0.003387
Coşkun [92]	0.003448
Chen [93]	0.003466
Akhtar [94]	0.003467
Akhtar [95]	0.003541

Appendix A.1.3 Data Clustering Using Gephi

One of the objectives of this study was to determine a series of categories that classify articles based on common elements. As Gephi has functionalities for the detection of clusters or communities in data samples, it was considered appropriate to use it to directly establish the categories, or as a guide when establishing them. The process carried out in the article by Mishra et al. [19] was taken as a reference.

Gephi uses the Louvain method to generate clusters, that is, groups of articles sharing the most citations with each other. The Louvain method obtains better results than other modularity methods in terms of computing time, which allows it to analyse very large networks. In addition, it achieves very good results as shown by contrast with networks of already known communities. Another of its aspects of interest is that it allows different degrees of resolution depending on the level of localism

or the scale of the community that is to be detected. This is a very attractive aspect if the intention is to establish integrated micro-communities within others of a larger scale. The method is applied in two phases. First, it establishes small communities by optimizing modularity in a local way. It then adds the nodes of the same community and builds new networks whose nodes represent communities. This process is repeated iteratively until a maximum of modularity is reached [96].

Applying this method to the list of articles in this study resulted in three clusters that are represented in Figure A7. Then, the contents and research areas of all the papers in each cluster were carefully analysed in an attempt to discover common patterns and elements that justified why they were associated in the same group. It was found that studies in cluster 1 were concerned with the application of Industry 4.0 to sustainable supply chains without using Big Data analytics, a technological element that was predominant in the other two clusters. However, with respect to modules 2 and 3, it was impossible to find a clear pattern that accounted for the difference between the two clusters. Due to this, the definition of the clusters proposed by Gephi was discarded.

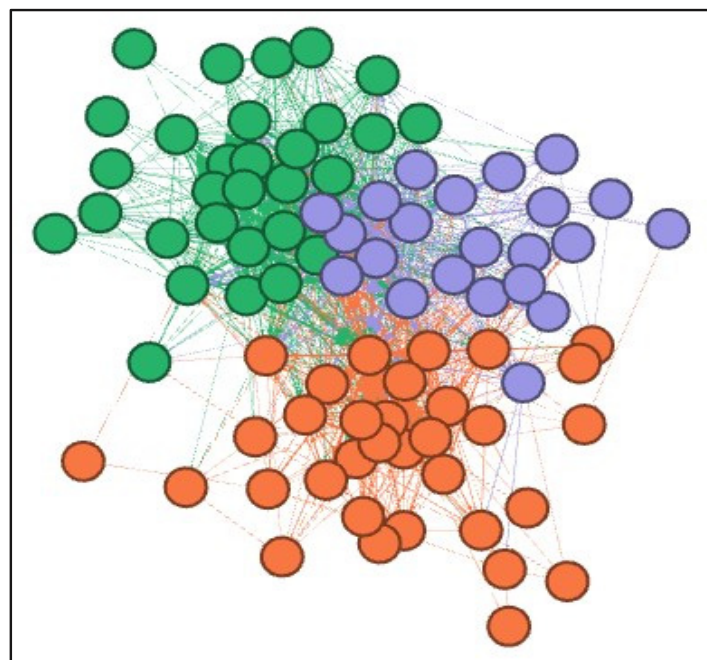


Figure A7. Structure of the three clusters.

References

1. Ortiz, O.; Castells, F.; Sonnemann, G. Sustainability in the construction industry: A review of recent developments based on LCA. *Constr. Build. Mater.* **2009**, *23*, 28–39. [\[CrossRef\]](#)
2. Maes, M.J.A.; Jones, G.; Toledano, M.B.; Milligan, B. Mapping synergies and trade-offs between urban ecosystems and the sustainable development goals. *Environ. Sci. Policy* **2019**, *93*, 181–188. [\[CrossRef\]](#)
3. Azevedo, S.G.; Cruz-Machado, V.; Carvalho, H.; Duarte, S. Influence of Green and Lean Upstream Supply Chain Management Practices on Business Sustainability. *IEEE Trans. Eng. Manag.* **2012**, *59*, 753–765. [\[CrossRef\]](#)
4. Srivastava, S.M. Green supply-chain management: A state-of-the-art literature review. *Int. J. Manag. Rev.* **2007**, *9*, 53–80. [\[CrossRef\]](#)
5. Pagell, M.; Shevchenko, A. Why Research in Sustainable Supply Chain Management Should Have no Future. *J. Supply Chain Manag.* **2014**, *50*, 44–55. [\[CrossRef\]](#)
6. Luthra, S.; Mangla, S.K. Evaluating challenges to Industry 4.0 initiatives for supply chain sustainability in emerging economies. *Process. Saf. Environ. Prot.* **2018**, *117*, 168–179. [\[CrossRef\]](#)
7. Rejikumar, G.; Raja, V.; Arunprasad, P.; Jinil, P.; Sreeraj, K.M. Industry 4.0: Key findings and analysis from the literature arena. *Benchmarking Int. J.* **2019**, *26*, 2514–2542.

8. Stock, T.; Seliger, G. Opportunities of sustainable manufacturing in Industry 4.0. *Procedia CIRP* **2016**, *40*, 536–541. [[CrossRef](#)]
9. 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](#)]
10. Dubey, R.; Gunasekaran, A.; Childe, S.J.; Papadopoulos, T.; Helo, P. Supplier relationship management for circular economy. *Manag. Decis.* **2019**, *57*, 767–790. [[CrossRef](#)]
11. Witkowski, K. Internet of Things, Big Data, Industry 4.0 – Innovative Solutions in Logistics and Supply Chains Management. *Procedia Eng.* **2017**, *182*, 763–769. [[CrossRef](#)]
12. Jabbour, A.B.L.D.S.; Jabbour, C.J.C.; Filho, M.G.; Roubaud, D. Industry 4.0 and the circular economy: A proposed research agenda and original roadmap for sustainable operations. *Ann. Oper. Res.* **2018**, *270*, 273–286.
13. Hazen, B.; Skipper, J.B.; Ezell, J.D.; Boone, C.A. Big data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Comput. Ind. Eng.* **2016**, *101*, 592–598. [[CrossRef](#)]
14. Gosling, J.; Jia, F.; Gong, Y.; Brown, S. The role of supply chain leadership in the learning of sustainable practice: Toward an integrated framework. *J. Clean. Prod.* **2017**, *140*, 239–250. [[CrossRef](#)]
15. Kamble, S.S.; Gunasekaran, A.; Gawankar, S.A. Sustainable Industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives. *Process. Saf. Environ. Prot.* **2018**, *117*, 408–425. [[CrossRef](#)]
16. Mani, V.; Delgado, C.; Hazen, B.; Patel, P. Mitigating Supply Chain Risk via Sustainability Using Big Data Analytics: Evidence from the Manufacturing Supply Chain. *Sustainability* **2017**, *9*, 608. [[CrossRef](#)]
17. Rowley, J.; Slack, F. Conducting a literature review. *Manag. Res. News* **2004**, *27*, 31–39. [[CrossRef](#)]
18. Wohlin, C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, London, UK, 13–14 May 2014.
19. Mishra, D.; Gunasekaran, A.; Papadopoulos, T.; Childe, S.J. Big Data and supply chain management: A review and bibliometric analysis. *Ann. Oper. Res.* **2016**, *270*, 313–336. [[CrossRef](#)]
20. Collier, D. *The Comparative Method. Political Science: The State Of Discipline Ii*; Finifter, A.E., Ed.; American Political Science Association: Washington, DC, USA, 1993.
21. Mentzer, J.T.; DeWitt, W.; Keebler, J.S.; Min, S.; Nix, N.W.; Smith, C.D.; Zacharia, Z.G. DEFINING SUPPLY CHAIN MANAGEMENT. *J. Bus. Logist.* **2001**, *22*, 1–25. [[CrossRef](#)]
22. Tan, K.C.; Lyman, S.B.; Wisner, J.D. Supply chain management: A strategic perspective. *Int. J. Oper. Prod. Manag.* **2002**, *22*, 614–631.
23. Otto, A.; Kotzab, H. Does supply chain management really pay? Six perspectives to measure the performance of managing a supply chain. *Eur. J. Oper. Res.* **2003**, *144*, 306–320. [[CrossRef](#)]
24. Lummus, R.R.; Vokurka, R.J. Defining supply chain management: A historical perspective and practical guidelines. *Ind. Manag. Data Syst.* **1999**, *99*, 11–17. [[CrossRef](#)]
25. Gruchmann, T.; Seuring, S. Explaining logistics social responsibility from a dynamic capabilities perspective. *Int. J. Logist. Manag.* **2018**, *29*, 1255–1278. [[CrossRef](#)]
26. Rebs, T.; Brandenburg, M.; Seuring, S. System dynamics modeling for sustainable supply chain management: A literature review and systems thinking approach. *J. Clean. Prod.* **2019**, *208*, 1265–1280. [[CrossRef](#)]
27. Hussain, M.; Ajmal, M.; Gunasekaran, A.; Khan, M. Exploration of social sustainability in healthcare supply chain. *J. Clean. Prod.* **2018**, *203*, 977–989. [[CrossRef](#)]
28. Zimon, D.; Tyan, J.; Sroufe, R. DRIVERS OF SUSTAINABLE SUPPLY CHAIN MANAGEMENT: PRACTICES TO ALIGNMENT WITH UN SUSTAINABLE DEVELOPMENT GOALS. *Int. J. Qual. Res.* **2020**, *14*, 219–236. [[CrossRef](#)]
29. Zimon, D.; Tyan, J.; Sroufe, R. Implementing Sustainable Supply Chain Management: Reactive, Cooperative, and Dynamic Models. *Sustainability* **2019**, *11*, 7227. [[CrossRef](#)]
30. Gharaei, A.; Karimi, M.; Shekarabi, S.A.H. An integrated multi-product, multi-buyer supply chain under penalty, green, and quality control polices and a vendor managed inventory with consignment stock agreement: The outer approximation with equality relaxation and augmented penalty algorithm. *Appl. Math. Model.* **2019**, *69*, 223–254. [[CrossRef](#)]

31. Hao, Y.; Helo, P.; Shamsuzzoha, A. Virtual factory system design and implementation: Integrated sustainable manufacturing. *Int. J. Syst. Sci. Oper. Logist.* **2016**, *5*, 116–132. [[CrossRef](#)]
32. Izadikhah, M.; Saen, R.F. Assessing sustainability of supply chains by chance-constrained two-stage DEA model in the presence of undesirable factors. *Comput. Oper. Res.* **2018**, *100*, 343–367. [[CrossRef](#)]
33. Tajbakhsh, A.; Hassini, E. A data envelopment analysis approach to evaluate sustainability in supply chain networks. *J. Clean. Prod.* **2015**, *105*, 74–85. [[CrossRef](#)]
34. Ansari, Z.N.; Kant, R. Exploring the Framework Development Status for Sustainability in Supply Chain Management: A Systematic Literature Synthesis and Future Research Directions. *Bus. Strat. Environ.* **2017**, *26*, 873–892. [[CrossRef](#)]
35. Adolph, L. *German Standardization Roadmap: Industry 4.0, Version 2*; DKE Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE: Berlin, Germany, 2016.
36. Nagy, J.; Oláh, J.; Erdei, E.; Máté, D.; Popp, J. The role and impact of Industry 4.0 and the internet of things on the business strategy of the value chain—the case of Hungary. *Sustainability* **2018**, *10*, 3491. [[CrossRef](#)]
37. Plattform Industrie 4.0. *Industrie 4.0 Whitepaper FuE-Themen*; Plattform Industrie: Berlin, Germany, 2015.
38. Martínez-Olvera, C.; Mora-Vargas, J. A Comprehensive Framework for the Analysis of Industry 4.0 Value Domains. *Sustainability* **2019**, *11*, 2960. [[CrossRef](#)]
39. Birkel, H.S.; Veile, J.W.; Müller, J.M.; Hartmann, E.; Voigt, K.I. Development of a risk framework for Industry 4.0 in the context of sustainability for established manufacturers. *Sustainability* **2019**, *11*, 384. [[CrossRef](#)]
40. Lin, K.; Shyu, J.; Ding, K. A cross-strait comparison of innovation policy under industry 4.0 and sustainability development transition. *Sustainability* **2017**, *9*, 786. [[CrossRef](#)]
41. Thomas, A.; Haven-Tang, C.; Barton, R.; Mason-Jones, R.; Francis, M.; Byard, P. Smart Systems Implementation in UK Food Manufacturing Companies: A Sustainability Perspective. *Sustainability* **2018**, *10*, 4693. [[CrossRef](#)]
42. Varela, L.; Araújo, A.; Ávila, P.; Castro, H.; Putnik, G. Evaluation of the Relation between Lean Manufacturing, Industry 4.0, and Sustainability. *Sustainability* **2019**, *11*, 1439. [[CrossRef](#)]
43. Carvalho, N.G.P.; Chaim, O.; Cazarini, E.; Gerolamo, M. Manufacturing in the fourth industrial revolution: A positive prospect in Sustainable Manufacturing. *Procedia Manuf.* **2018**, *21*, 671–678. [[CrossRef](#)]
44. Flouris, I.; Giatrakos, N.; Deligiannakis, A.; Garofalakis, M.; Kamp, M.; Mock, M. Issues in complex event processing: Status and prospects in the Big Data era. *J. Syst. Softw.* **2017**, *127*, 217–236. [[CrossRef](#)]
45. Orenga-Rogla, S.; Chalmeta, R. Methodology for the Implementation of Knowledge Management Systems 2.0. *Bus. Inf. Syst. Eng.* **2017**, *61*, 195–213. [[CrossRef](#)]
46. Orenga-Rogla, S.; Chalmeta, R. Framework for implementing a big data ecosystem in organizations. *Commun. ACM* **2018**, *62*, 58–65. [[CrossRef](#)]
47. Shin, N.-H.; Choi, M.J. Ecological views of big data: Perspectives and issues. *Telemat. Inform.* **2015**, *32*, 311–320. [[CrossRef](#)]
48. Dubey, R.; Gunasekaran, A.; Papadopoulos, T.; Childe, S.J.; Shibin, K.T.; Wamba, S.F. Sustainable supply chain management: Framework and further research directions. *J. Clean. Prod.* **2017**, *142*, 1119–1130. [[CrossRef](#)]
49. De, A.; Wang, J.; Tiwari, M. Hybridizing Basic Variable Neighbourhood Search with Particle Swarm Optimization for Solving Sustainable Ship Routing and Bunker Management Problem. In Proceedings of the IEEE Transactions on Intelligent Transportation Systems, Auckland, New Zealand, 27–30 October 2019.
50. De, A.; Choudhary, A.; Tiwari, M.K. Multiobjective Approach for Sustainable Ship Routing and Scheduling With Draft Restrictions. *IEEE Trans. Eng. Manag.* **2017**, *66*, 35–51. [[CrossRef](#)]
51. Chen, L.; Jia, G. Environmental efficiency analysis of China's regional industry: A data envelopment analysis (DEA) based approach. *J. Clean. Prod.* **2017**, *142*, 846–853. [[CrossRef](#)]
52. Mongeon, P.; Paul-Hus, A. The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics* **2016**, *106*, 213–228. [[CrossRef](#)]
53. Brin, S.; Page, L. The anatomy of a large-scale hypertextual Web search engine. *Comput. Networks ISDN Syst.* **1998**, *30*, 107–117. [[CrossRef](#)]
54. Ma, N.; Guan, J.C.; Zhao, Y. Bringing PageRank to the citation analysis. *Inf. Process. Manag.* **2008**, *44*, 800–810. [[CrossRef](#)]
55. Shrouf, F.; Ordieres, J.; Miragliotta, G. Smart factories in Industry 4.0: A review of the concept and of energy Management Approached in Production Based on the Internet of Things paradigm. In Proceedings of the 2014 IEEE International Conference on Industrial Engineering and Engineering Management, Selangor Darul Ehsan, Malaysia, 9–12 December 2014; pp. 697–701.

56. Chae, B. (Kevin) Insights from hashtag #supplychain and Twitter Analytics: Considering Twitter and Twitter data for supply chain practice and research. *Int. J. Prod. Econ.* **2015**, *165*, 247–259.
57. Dubey, R.; Gunasekaran, A.; Childe, S.J.; Wamba, S.F.; Papadopoulos, T. The impact of big data on world-class sustainable manufacturing. *Int. J. Adv. Manuf. Technol.* **2015**, *84*, 631–645. [[CrossRef](#)]
58. Zhang, Y.; Ren, S.; Liu, Y.; Si, S. A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. *J. Clean. Prod.* **2017**, *142*, 626–641. [[CrossRef](#)]
59. Papadopoulos, T.; Gunasekaran, A.; Dubey, R.; Altay, N.; Childe, S.J.; Wamba, S.F. The role of Big Data in explaining disaster resilience in supply chains for sustainability. *J. Clean. Prod.* **2017**, *142*, 1108–1118. [[CrossRef](#)]
60. Zhao, R.; Liu, Y.; Zhang, N.; Huang, T. An optimization model for green supply chain management by using a big data analytic approach. *J. Clean. Prod.* **2017**, *142*, 1085–1097. [[CrossRef](#)]
61. Wu, K.-J.; Liao, C.-J.; Tseng, M.-L.; Lim, M.K.; Hu, J.; Tan, K.H. Toward sustainability: Using big data to explore the decisive attributes of supply chain risks and uncertainties. *J. Clean. Prod.* **2017**, *142*, 663–676. [[CrossRef](#)]
62. Fawcett, S.E.; Waller, M. Supply Chain Game Changers-Mega, Nano, and Virtual Trends-And Forces That Impede Supply Chain Design (i.e., Building a Winning Team). *J. Bus. Logist.* **2014**, *35*, 157–164. [[CrossRef](#)]
63. Rehman, M.H.U.; Chang, V.; Batool, A.; Wah, T.Y. Big data reduction framework for value creation in sustainable enterprises. *Int. J. Inf. Manag.* **2016**, *36*, 917–928. [[CrossRef](#)]
64. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572. [[CrossRef](#)]
65. Smith, K.M.; Crookes, E.; Crookes, P.A. Measuring research 'impact' for academic promotion: Issues from the literature. *J. High. Educ. Policy Manag.* **2013**, *35*, 410–420. [[CrossRef](#)]
66. Sanders, A.; Elangeswaran, C.; Wulfsberg, J. Industry 4.0 implies lean manufacturing: Research activities in industry 4.0 function as enablers for lean manufacturing. *J. Ind. Eng. Manag.* **2016**, *9*, 811. [[CrossRef](#)]
67. Song, M.; Cen, L.; Zheng, Z.; Fisher, R.; Liang, X.; Wang, Y.; Huisingh, D. How would big data support societal development and environmental sustainability? Insights and practices. *J. Clean. Prod.* **2017**, *142*, 489–500. [[CrossRef](#)]
68. Kusiak, A. Smart manufacturing. *Int. J. Prod. Res.* **2017**, *56*, 508–517. [[CrossRef](#)]
69. Cerchione, R.; Esposito, E. A systematic review of supply chain knowledge management research: State of the art and research opportunities. *Int. J. Prod. Econ.* **2016**, *182*, 276–292. [[CrossRef](#)]
70. Tiwari, S.; Wee, H.; Daryanto, Y. Big data analytics in supply chain management between 2010 and 2016: Insights to industries. *Comput. Ind. Eng.* **2018**, *115*, 319–330. [[CrossRef](#)]
71. Kuo, T.C.; Smith, S. A systematic review of technologies involving eco-innovation for enterprises moving towards sustainability. *J. Clean. Prod.* **2018**, *192*, 207–220. [[CrossRef](#)]
72. Zhong, R.Y.; Xu, C.; Chen, C.; Huang, G.Q. Big Data Analytics for Physical Internet-based intelligent manufacturing shop floors. *Int. J. Prod. Res.* **2015**, *55*, 2610–2621. [[CrossRef](#)]
73. Rodriguez, L.; Da Cunha, C. Impacts of Big Data analytics and absorptive capacity on sustainable supply chain innovation: A conceptual framework. *LogForum.* **2018**, *14*, 151–161. [[CrossRef](#)]
74. Hopkins, J.; Hawking, P. Big Data Analytics and IoT in logistics: A case study. *Int. J. Logist. Manag.* **2018**, *29*, 575–591. [[CrossRef](#)]
75. Richey, R.G.; Morgan, T.R.; Lindsey-Hall, K.; Adams, F.G. A global exploration of Big Data in the supply chain. *Int. J. Phys. Distrib. Logist. Manag.* **2016**, *46*, 710–739. [[CrossRef](#)]
76. Jensen, J.; Remmen, A. Enabling Circular Economy Through Product Stewardship. *Procedia Manuf.* **2017**, *8*, 377–384. [[CrossRef](#)]
77. Liu, W.-K.; Yen, C.-C. Optimizing Bus Passenger Complaint Service through Big Data Analysis: Systematized Analysis for Improved Public Sector Management. *Sustainability* **2016**, *8*, 1319. [[CrossRef](#)]
78. Keivanpour, S.; Kadi, D.A. Perspectives for application of the internet of things and Big Data analytics on end of life aircraft treatment. *Int. J. Sustain. Aviat.* **2018**, *4*, 202–220. [[CrossRef](#)]
79. Hazen, B.; Boone, C.A.; Wang, Y.; Khor, K.S. Perceived quality of remanufactured products: Construct and measure development. *J. Clean. Prod.* **2017**, *142*, 716–726. [[CrossRef](#)]
80. Badiezhadeh, T.; Saen, R.F.; Samavati, T. Assessing sustainability of supply chains by double frontier network DEA: A big data approach. *Comput. Oper. Res.* **2018**, *98*, 284–290. [[CrossRef](#)]

81. How, B.S.; Lam, H.L.; Shen, H.B.; Loong, L.H. Sustainability evaluation for biomass supply chain synthesis: Novel principal component analysis (PCA) aided optimisation approach. *J. Clean. Prod.* **2018**, *189*, 941–961. [[CrossRef](#)]
82. Bányai, T.; Illés, B.; Bányai, Á. Smart Scheduling: An Integrated First Mile and Last Mile Supply Approach. *Complexity* **2018**, *2018*, 1–15.
83. Malesios, C.; Psarakis, S. Comparison of the h-index for different fields of research using bootstrap methodology. *Qual. Quant.* **2012**, *48*, 521–545. [[CrossRef](#)]
84. Gani, A.; Siddiqa, A.; Shamshirband, S.; Hanum, F. A survey on indexing techniques for big data: Taxonomy and performance evaluation. *Knowl. Inf. Syst.* **2015**, *46*, 241–284. [[CrossRef](#)]
85. Svensson, G. Aspects of sustainable supply chain management (SSCM): Conceptual framework and empirical example. *Supply Chain Manag. Int. J.* **2007**, *12*, 262–266. [[CrossRef](#)]
86. Alfian, G.; Syafrudin, M.; Rhee, J. Real-Time Monitoring System Using Smartphone-Based Sensors and NoSQL Database for Perishable Supply Chain. *Sustainability* **2017**, *9*, 2073. [[CrossRef](#)]
87. Ahearn, M.C.; Armbruster, W.; Young, R. Big Data's Potential to Improve Food Supply Chain Environmental Sustainability and Food Safety. *Int. Food Agribus. Manag. Rev.* **2016**, *19*, 155–171.
88. Akhtar, P.; Khan, Z.; Frynas, J.G.; Tse, Y.K.; Rao-Nicholson, R. Essential Micro-foundations for Contemporary Business Operations: Top Management Tangible Competencies, Relationship-based Business Networks and Environmental Sustainability. *Br. J. Manag.* **2017**, *29*, 43–62. [[CrossRef](#)]
89. Ardito, L.; Petruzzelli, A.M.; Panniello, U.; Garavelli, A.C. Towards Industry 4.0: Mapping digital technologies for supply chain management-marketing integration. *Bus. Process Manag. J.* **2019**, *25*, 323–346. [[CrossRef](#)]
90. Coble, K.H.; Mishra, A.K.; Ferrell, S.; Griffin, T. Big Data in Agriculture: A Challenge for the Future. *Appl. Econ. Perspect. Policy* **2018**, *40*, 79–96. [[CrossRef](#)]
91. Ji, S.; Sun, Q. Low-Carbon Planning and Design in B&R Logistics Service: A Case Study of an E-Commerce Big Data Platform in China. *Sustainability* **2017**, *9*, 2052.
92. Coşkun, S.; Polat, O.; Özgür, L. Analytic hierarchy process and data envelopment analysis methods based model proposal for improving supplier efficiencies and a case study. *J. Mehmet Akif Ersoy Univ. Econ. Adm. Sci. Fac.* **2019**, 11672.
93. Chen, R.-Y. Intelligent IoT-Enabled System in Green Supply Chain using Integrated FCM Method. *Int. J. Bus. Anal.* **2015**, *2*, 47–66. [[CrossRef](#)]
94. Akhtar, P.; Khan, Z.; Rao-Nicholson, R.; Zhang, M. Building relationship innovation in global collaborative partnerships: Big data analytics and traditional organizational powers. *R&D Manag.* **2016**, *49*, 7–20.
95. Akhtar, P.; Tse, Y.K.; Khan, Z.; Rao-Nicholson, R. Data-driven and adaptive leadership contributing to sustainability: Global agri-food supply chains connected with emerging markets. *Int. J. Prod. Econ.* **2016**, *181*, 392–401. [[CrossRef](#)]
96. Blondel, V.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, É. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P10008. [[CrossRef](#)]

