



Article

# Big Data in Education. A Bibliometric Review

José-Antonio Marín-Marín <sup>1</sup>, Jesús López-Belmonte <sup>2</sup>, Juan-Miguel Fernández-Campoy <sup>1</sup> and José-María Romero-Rodríguez <sup>1,\*</sup>

<sup>1</sup> Department of Didactics and School Organization, University of Granada, 18071 Granada, Spain

<sup>2</sup> International University of Valencia, 46002 Valencia, Spain

\* Correspondence: romejo@ugr.es

Received: 4 June 2019; Accepted: 17 July 2019; Published: 25 July 2019



**Abstract:** The handling of a large amount of data to analyze certain behaviors is reaching a great popularity in the decade 2010–2020. This phenomenon has been called Big Data. In the field of education, the analysis of this large amount of data, generated to a greater extent by students, has begun to be introduced in order to improve the teaching–learning process. In this paper, it was proposed as an objective to analyze the scientific production on Big Data in education in the databases Web of Science (WOS), Scopus, ERIC, and PsycINFO. A bibliometric study was carried out on a sample of 1491 scientific documents. Among the results, the increase in publications in 2017 and the configuration of certain journals, countries and authors as references in the subject matter stand out. Finally, potential explanations for the study findings and suggestions for future research are discussed.

**Keywords:** Big Data; education; bibliometric study; Internet

## 1. Introduction

Big Data is a concept that is currently in fashion and has been in specialized literature for more than a decade, alluding to the large amount of data that is generated at every moment as a result of technological evolution and the interactions of people in digital spaces (Waller and Fawcett 2013). However, it is only recently that it has had its greatest apogee and impact as an object of research as a result of technological advances and the development of platforms for interaction between users and these with the content, leading to an enormous amount of data (Ghani et al.). Specifically, Big Data refers to the large volume of data generated because of the development of technology and the continuous actions and interactions of users in digital environments (Hussain and Cambria 2018). Other concepts related to Big Data are data learning mining or learning analytics. Data learning mining is all those techniques and procedures used to extract useful and relevant information from the large amount of data reported from educational platforms (Menon et al. 2017). On the other hand, learning analytics is a construct that is derived from data mining and alludes to the management, processing, and analysis of students' educational data, which are studied with the purpose of improving and optimizing the learning process (Liang et al. 2016).

That is why, today, society is in what experts call the Big Data era, promulgating new challenges and benefits through the analysis of all data generated in environments characterized by high quantification (Pugna et al. 2019).

Since the arrival of the new millennium, services such as the Internet and the development of the Web began to record data from users, their movements and interactions, creating a large bank of useful and relevant information, whose analysis reports great potentialities to study the needs and demands of people (Chen et al. 2012; Khan et al. 2018).

Technological development and the emergence of popular social networks have led people to become active agents in digital media, exponentially multiplying the amount of data generated (Ni et al. 2016).

All this has led to a great interest on the part of researchers in studying all aspects concerning the enormous presence of data in all aspects of people's lives (Williamson 2015; Williams et al. 2017). Thus, the European Commission stated that the Horizon 2020 report would be a major step towards the study of Big Data, with the aim of developing strategies to conduct research and innovation in this field of knowledge (Jin et al. 2015).

The purpose of Big Data analysis is to collect a set of data from various electronic sources to be transformed into relevant information in order to improve the services which the user habitually accesses (Jagadish 2016).

Big Data is nourished by an era marked by the connectivity of people (Veltri 2017), where the action of creating contents, sharing and interacting with the rest of users in the community are the order of the day (Hussain and Cambria 2018). This provides a great opportunity to know—in addition to the needs—the psychological state of people and their behaviour in virtual spaces (Eichstaedt et al. 2015).

Given the peculiarities of the society in which we live, the data are growing at great speed (Al Nuaimi et al. 2015). So much so, that volume, speed, variety, veracity, and value are already spoken of as fundamental characteristics of the data and that are inherent to Big Data. They present a disorganized structure and are in various formats such as text, image, voice, and video (Injadat et al. 2016).

In order to analyse all the data in the digital environment, the concept of data science arises with the intention of managing and interpreting each and every one of the data by means of specialised programmes with high processing capacity (Hicks and Irizarry 2018). These developments have led to the evolution of predictive analytics (Waller and Fawcett 2013), to adapt services to current trends demanded by the user (Saiki et al. 2018). Therefore, the data are used to predict and make decisions about the future (Ghani et al.), based on a strategic design that analyzes the requirements of the audience (Perlado-Lamo-de-Espinosa et al. 2019).

According to Moreno-Carriles (2018), the literature reveals that the treatment of Big Data has expanded into different fields of action, such as security, customer service, public services, preservation of the environment, the economy, finance, in addition to education, which is the field that interests us in this study.

The Big Data that has mainly been exploited in the business world today is already being widely used in education (Aretio 2017), finding us in a new phase of teaching and learning based on the study of data generated by students (Gibson 2017).

All the data derived from the different educational agents (teachers and learners) are currently being processed in order to improve the quality and experience of learning processes in digital environments (Liang et al. 2016).

Likewise, the data source produced by educational content management platforms is being used to develop tools and services adapted to the singularities of contemporary education, highly conditioned by the development of educational technology (Merceron et al. 2015). The immersion of the students in a distance and ubiquitous education has caused a great flow of data about their developed activity (Seufert et al. 2019).

However, experts such as Menon et al. (2017) consider that data mining techniques in the field of education—to this day—are not completely successful, so not all meaningful and valuable information is extracted. This is due to the fact that the handling and treatment of Big Data requires the collaboration of teachers with specialists, with the objective of being able to obtain the relevant information from the data reported by the use of tools and digital resources of an educational nature (Huda et al. 2017). This allows learners to perform all kinds of actions in virtual spaces, whose generated data are used to obtain knowledge about their activity, performance and satisfaction (Elia et al. 2019).

An effective analysis of Big Data contributes to the promotion of new and better educational experiences (Reidenberg and Schaub 2018), to an improvement of didactic programming tasks on the part of teachers with the help of scientists specializing in data analysis, to an efficient selection of strategies and decision making to approach the formative process, adequate to the demands of a

learning group increasingly familiar with technology, seeking innovative learning as a result of the study of data (Huda et al. 2018), and all of this based on a predictive analysis of the data collected (Daniel 2015; Daniel 2017).

Therefore, Big Data and analytics of the interactions of educational agents in virtual environments are positioned as new ways to solve the shortcomings of the educational system (Picciano 2012), in such a way as to improve productivity, innovation (Sanchez and Ball 2015), and the personalisation of learning (Dishon 2017). As a result, it was proposed as an objective to analyze the scientific output, understood as the published articles on Big Data in education in the Web of Science (WOS), Scopus, ERIC, and PsycINFO databases. Consequently, the following research questions were identified:

RQ1. What is the state of scientific production over time?

RQ2. Which journals and countries concentrate the greatest scientific production on Big Data in education?

RQ3. Which are the articles of greater impact in the area of Big Data in education?

RQ4. What are the main lines of research in this field that are derived from the keywords of scientific articles?

## 2. Method

This study is characterized by following a bibliometric analysis methodology (Glänzel and Schoepflin 1999). So, following the guidelines and criteria of bibliometrics (Ardanuy Baró 2012), was first established the combination of keywords: “Big Data” AND education. This combination was introduced in the search engine of the different databases. Thus, the scientific production is collected, in article format, from 2010 to 2018. The search took place during the second quarter of 2019, so all indexed literature is included in the year 2018.

### 2.1. Sample

The unit of analysis was composed by the scientific articles indexed in WOS, Scopus, ERIC, and PsycINFO that included in the title, abstract, or keywords the terms Big Data and education. Finally, the sample consisted of journal articles that met the inclusion and exclusion criteria. Inclusion criteria were considered: (i) scientific articles published in journals and peer-reviewed; (ii) year of publication since the term appears in the literature until 2018; (iii) search descriptors appear in the title, abstract or keywords; (iv) published in English language. Instead, the exclusion criteria were: (i) documents not subject to exhaustive peer review (reviews, theses, books, book chapters, conference proceedings, or technical reports); (ii) articles that did not belong to the limited time period; (iii) descriptors are not included in the title, abstract, or keywords; (iv) the language of publication is not English.

From its application, the sample of analysis was composed into 1491 documents: 491 in WOS; 706 in Scopus; 174 in ERIC; and 120 in PsycINFO.

### 2.2. Data Analysis

Data analysis was performed from information extracted from the four databases. Excel and VOSviewer version 1.6.7 (Centre for Science and Technology, Leiden University, Leiden, The Netherlands) programs were used to support the analysis and graphical representation of the data.

On the other hand, the analysis variables were established from the review of previous bibliometric studies in the area of social sciences, with a topic similar to the object of study (Batanero et al. 2019; Hinojo-Lucena et al. 2019; Oravec et al. 2019; Rodríguez-García et al. 2019; Sudolska et al. 2019):

Publications by year.

Journals and countries with the highest number of articles.

Articles with greater impact.

Keywords.

In addition, the bibliometric laws of Price and Bradford were applied to verify diachronic productivity, i.e., productivity over the years (Price 1986) and to establish the nucleus formed by the journals with the largest number of articles (Urbizagástegui Alvarado 2016).

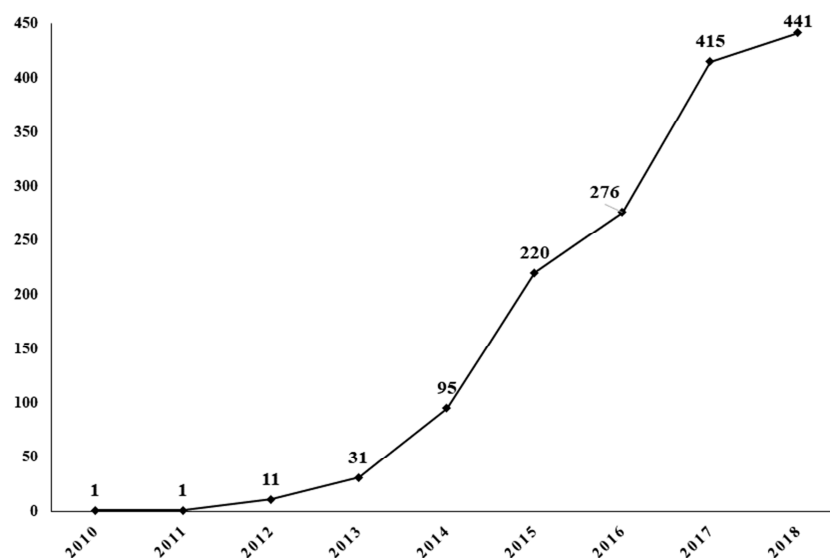
### 3. Results

The publications per year are mostly concentrated between 2016 and 2018, covering 75.92% of the articles published on Big Data in education. Likewise, its origin in literature begins in 2010, although the flow of publications begins in 2012 (Table 1).

**Table 1.** Publication per year of scientific articles on the databases.

Year	Articles Number			
	WOS	Scopus	ERIC	PsycINFO
2010	0	1	0	0
2011	0	0	0	1
2012	2	6	3	0
2013	6	13	9	3
2014	23	46	20	6
2015	75	98	27	20
2016	92	116	37	31
2017	146	200	38	31
2018	147	226	40	28
Total	491	706	174	120

Price’s law establishes that, after 10 years, the scientific literature tends to double, at the same time that it fixes three stages in the development and consolidation of the literature (Price 1986). In this respect, the premise of the duplicity of literature is confirmed, since by the year 2018 literature is much higher than in 2010 (Figure 1). Looking at the graph, the development stages of Price are set between 2010 and 2012 (precursors stage) and from 2012 to the present day, in the exponential growth stage.



**Figure 1.** Diachronic output on Big Data in education.

On the other hand, Bradford’s law indicates that, by making an equitable distribution of articles by zones, a small cluster of journals (centre) is formed which collect an equivalent quantity of articles to the rest of zones (Urbizagástegui Alvarado 2016). This is confirmed in the literature published on Big Data in education, where a small group of journals are the ones that collect the most articles

on this topic (Figure 2). Specifically, in WOS the total is 346 journals and 491 articles, distributed in five areas with the same number of articles approximately ( $M = 98.2$ ). In this sense, it is observed that the nucleus conformed by 16 journals contains a similar amount of documents to the rest of the zones. In Scopus, 419 journals and 706 articles are collected, distributed in five other zones ( $M = 141.2$ ); the centre consists of six journals. In ERIC, there are 112 journals and 174 articles, grouped into five zones ( $M = 34.8$ ) and with a core made up of four journals. Finally, PsycINFO groups 83 journals and 120 articles into five zones ( $M = 24$ ) and with a core of five journals.

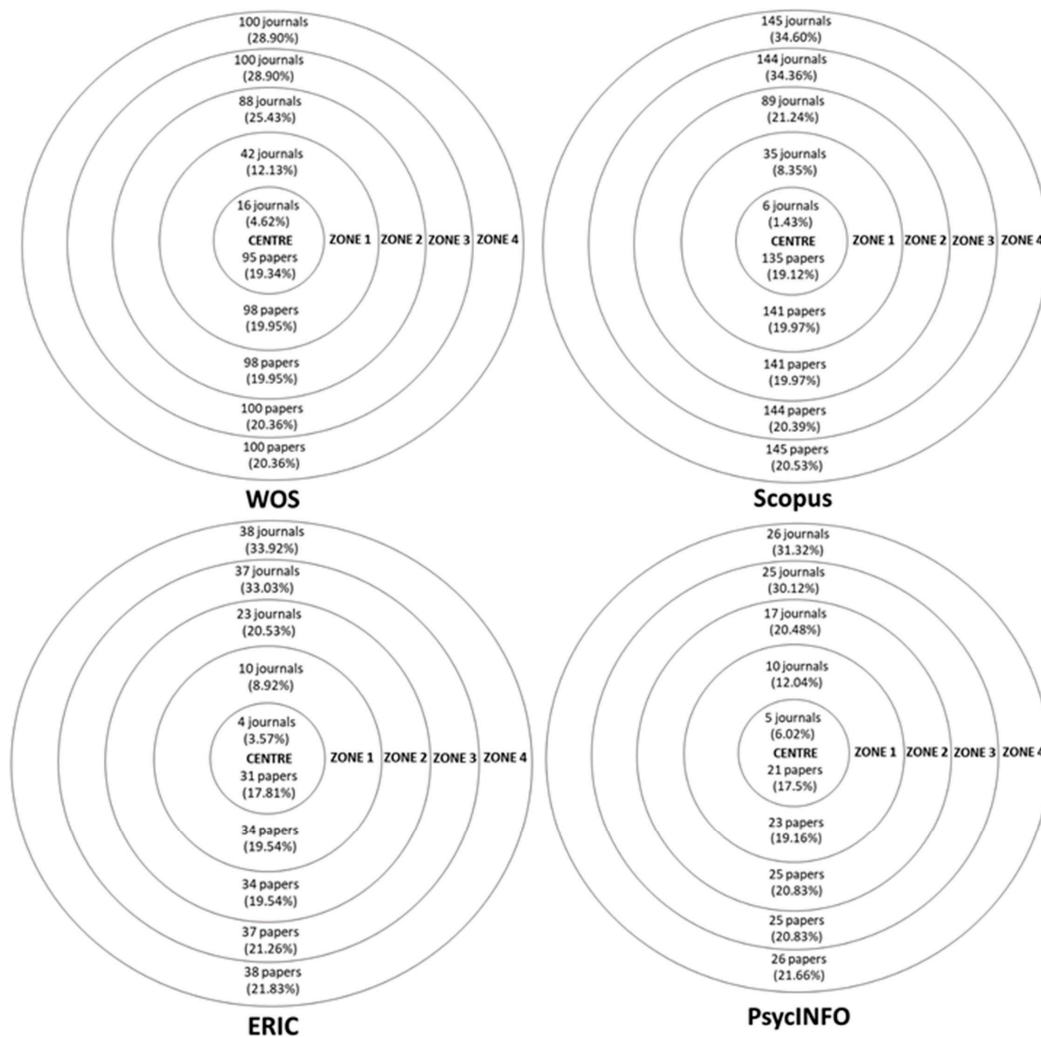


Figure 2. Bradford’s scattering zone of scientific journals on Big Data in education.

These journals, which make up the core, have a much higher than average number of articles. Among them, coinciding in the core of WOS and Scopus: Agro Food Industry and International Journal of Emerging Technologies in Learning (iJET). In WOS and ERIC: Theory and Research in Education. Moreover, in WOS and PsycINFO: Behaviour & Information Technology (BIT) (Table 2).

**Table 2.** Journals that make up the centre in WOS, Scopus, ERIC and PsycINFO.

Journals			
WOS	Scopus	ERIC	PsycINFO
ESTP	JAOT	ISEDJ	Computers in Human Behavior
Agro Food Industry Hi Tech	Technical Bulletin	Journal of Learning Analytics	Neurocomputing
ijET	Agro Food Industry Hi Tech	JISE	DSJIE
Engineering	Frontiers of Computer Science	Theory and Research in Education	Learning, Media and Technology
Big Data	KUYEB		BIT
Theory and Research in Education	ijET		
IJACSA			
Big Data society			
IEEE access			
Medical Teacher			
American Statistician			
BIT			
EJMSTE			
LNET			
Sustainability			
Technology Knowledge and Learning			

Note: ESTP = Educational Sciences Theory Practice; IJACSA = International Journal of Advanced Computer Science and Applications; EJMSTE = Eurasia Journal of Mathematics Science and Technology Education; JAOT = Journal of Advanced Oxidation Technologies; ISEDJ = Information Systems Education Journal; KUYEB = Kuram Ve Uygulamada Egitim Bilimleri; DSJIE = Decision Sciences Journal of Innovative Education; JISE = Journal of Information Systems Education; LNET = Lecture notes in Educational Technology.

In relation to the countries with the highest scientific output, the top 10 are collected in each database. The United States stands out above all others, being the country with the largest amount of documents (30.65% of total publications on Big Data in education). China presents the second-largest collection of articles (21.66%) and the United Kingdom is in the third position (10.26%). Below are Australia (4.35%), Canada (3.62 %), Germany (2.34%), India (2.74%), Italy (1.81%), Sweden (1.40%), Saudi Arabia (1.20%), South Korea (2.41%), Japan (1.34%), and Brazil (0.60%) (Table 3).

As for the articles with the greatest impact, depending on the number of citations, it was taken as a criterion that they had more than 100 citations. Thus, six articles are established that met this criterion (Table 4). The first of these presents 2921 citations in WOS and Scopus, entitled “Business intelligence and analytics: From Big data to big impact”, in which the authors analyse how big data influences society and specifically business, challenges, and opportunities associated with the company research and education are identified (Chen et al. 2012). Behind it is “Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management”, a study conducted on how Big data helps improve supply chain management, they show that these terms are relevant to supply chain research and education (Waller and Fawcett 2013), with a sum of 562 citations. “Psychological Language on Twitter Predicts County-Level Heart Disease Mortality”, the authors gather analysis of Twitter data to predict cardiovascular disease, the education is significant in relation to big data analysis to predict such pathology (Eichstaedt et al. 2015), with 332 citations.

“Applications of big data to smart cities”, a study is collected that applies big data to improve urban services, highlighting improvements in the public education service (Al Nuaimi et al. 2015), with 230 citations. “Big Data and analytics in higher education: Opportunities and challenges”, the authors carry out an analysis of the advantages and challenges of applying big data in university education (Daniel 2015), with 179 citations. “The evolution of big data and learning analytics in American higher education”, a study is carried out that gathers the advances in analytical data technology in higher education in America (Picciano 2012), with 125 citations.

**Table 3.** Countries with the highest scientific production in WOS, Scopus, ERIC, and PsycINFO.

Country	Database			
	WOS	Scopus	ERIC	PsycINFO
United States (USA)	188	238	18	13
China	84	223	12	4
United Kingdom (UK)	70	71	8	4
Australia	32	27	5	1
Canada	19	30	4	1
Germany	19	15	1	0
India	18	20	2	1
Italy	13	13	1	0
Sweden	11	9	1	0
Saudi Arabia	9	8	1	0
South Korea	8	27	1	0
Japan	6	13	1	0
Brazil	3	4	2	0

**Table 4.** The most cited articles in WOS, Scopus, ERIC, and PsycINFO.

Reference	Year	Citations			
		WOS	Scopus	ERIC	PsycINFO
Chen et al. (2012)	2012	1111	1810	0	0
Waller and Fawcett (2013)	2013	242	320	0	0
Eichstaedt et al. (2015)	2015	111	142	0	79
Al Nuaimi et al. (2015)	2015	92	138	0	0
Daniel (2015)	2015	67	112	0	0
Picciano (2012)	2012	0	125	0	0

Finally, the networks map between keywords reflects the relations generated between them (Figure 3). The size of the words indicates their frequency of appearance and a greater amount of connections with other descriptors. There are also three distinct clusters, each with a different colour (red, blue, and green). The red cluster is led by the concept “approach”, with the descriptors “analytic”, “science”, “article”, and “technique” prevailing. The blue cluster is headed by “platform”, highlighting the keywords “innovation”, “service”, “person”, “factor”, “health”, and “industry”. Finally, in the green cluster, “student” stands out and includes descriptors linked mainly to education: “teaching”, “university”, “problem”, “content”, “attention”, “training”, and “algorithm”.





As an answer to the first of the questions posed in this study (What is the state of scientific production over time?), it should be noted that this is a young phenomenon and, therefore, one that has only recently come into being. This is demonstrated by the fact that the first scientific publications related to the subject do not begin to see the light until 2010. Although, it is no less true that since 2012 and up to the present time they have increased exponentially, as a result of the boom that this phenomenon has been experienced in the field of business, social networks and education (Bennett 2015).

Most of the research related to the study of Big Data (the second question of the study presented here) is concentrated, as far as the publication and dissemination of scientific results are concerned, in very specific journals located in countries or environments with a clear English or Anglo-Saxon tradition. This casual circumstance is related to the fact that these are some of the major environments in which these new currents of thought. In these environments, they are more widely developed, rooted, and consolidated, to the point that, in recent years, they have begun to become an outstanding element of dissemination to the rest of the developed countries of all these new approaches in the treatment and management of data, as elements of clear individualization, improvement, and efficiency of the teaching–learning processes (Caballero 2013).

By countries, and at a high level of agreement with the ideas outlined in the preceding paragraph, it should be noted that the United States is the country with the highest production, followed by China, the United Kingdom, and Canada, as far as scientific publications related to Big Data are concerned. Once again, there is evidence of the progressive development that trend of technological thought related to Big Data has been experiencing in Anglo-Saxon countries, to the point that they have become the main cultural window for the development of all these tools, especially in the field of education and, more specifically, the teaching–learning processes (DatAnalysis 15M 2013).

The only exception, with respect to the hegemonic countries in the use and disclosure of Big Data, is China, which appears in second place. This fact, although it goes a little out of the basic pattern, because it is not a country of Anglo-Saxon culture or English-speaking, is not surprising. It is well known that China is configured as a leading and technologically advanced nation that even becomes a pioneer in the development and implementation of many of the most important technological advances that end up reaching the main developed countries, including the United States itself (Medici 2009).

Although Germany does not end up occupying a predominant role with respect to the use, handling, and expansion of the technological tools assigned to Big Data, it is configured as the great gateway to Europe of the main technological advances related to the treatment of the complex data and information chains. As in the case of China, this result is not surprising either, because, with regard to the European continent, Germany represents the flagship of economic and technological prosperity. It is well above most of the countries that make up the European Union, and therefore ends up becoming the main introducer and engine of all kinds of advances, as well as a clear model to imitate (Hernández 2012).

With regard to the articles with the greatest impact linked to Big Data (the third question in the study), those that deal with topics clearly related to the field of business and production processes stand out, followed by those linked to the health field, in particular to the improvement of people's health levels and quality of life in elderly or elderly individuals. However, recent studies try to analyze the benefits of Big Data as a tool for the collection of data to analyze the development, design, and implementation of teaching–learning processes in their different phases and levels are gaining much prominence. With the idea of providing them with greater quality, efficiency, and significance, as well as to articulate the means and strategies that make it possible to individualize them and adapt them to the characteristics, needs, and interests of the students, in this way, we guarantee that each student, during their development, receives all that they need and that, therefore, we offer them the possibility of carrying out the whole teaching–learning process with high doses of efficiency and quality, contributing to the significance of the learning (Área 2011; Salazar 2016).

Ultimately, the main lines of research linked to the phenomenon of Big Data (fourth question of the study) show an almost absolute coincidence with the most important topics that have been working scientific articles of greater impact. This is evidenced by the fact that some of the main lines of research regarding Big Data are those that are found as a central topic in clusters. However, it also appears as a promising and very current line of research, which focuses on the figure of the student to place special emphasis on the knowledge of all those methodologies and strategies of a didactic nature. It has been insisting on the convenience and the need to evaluate the way in which the teaching–learning processes are being developed for the articulation, design, and implementation of perfectly individualized intervention procedures adapted to the needs of the student (Dussel 2014; Martín-Barbero 2012).

**Author Contributions:** Conceptualization, J.L.-B.; methodology, J.-M.R.-R. and J.-A.M.-M.; software, J.-M.R.-R.; formal analysis, J.-A.M.-M.; investigation, J.L.-B. and J.-M.R.-R.; resources, J.-M.F.-C.; data curation, J.L.-B.; writing—original draft preparation, J.-A.M.-M., J.-M.F.-C. and J.L.-B.; writing—review and editing, J.-M.R.-R. and J.-M.F.-C.; visualization, J.-M.R.-R.; supervision, J.-A.M.-M.

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

**Acknowledgments:** To the researchers of the research group AREA (HUM-672). Research group by belonging to the Ministry of Education and Science of the Junta de Andalucía and based in the Department of Didactics and School Organization of the Faculty of Education Sciences of the University of Granada.

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

## References

- Al Nuaimi, Eiman, Hind Al Neyadi, Nader Mohamed, and Jameela Al-Jaroodi. 2015. Applications of big data to smart cities. *Journal of Internet Services and Applications* 6: 1–15. [CrossRef]
- Ardanuy Baró, Jordi. 2012. *Breve Introducción a la Bibliometría*. Barcelona: University of Barcelona.
- Área, Manuel. 2011. Tic, identidad digital y educación. *Reencuentro* 62: 97–99.
- Aretio, Lorenzo García. 2017. Educación a distancia y virtual: calidad, disrupción, aprendizajes adaptativo y móvil. *RIED. Revista Iberoamericana de Educación a Distancia* 20: 9–25. [CrossRef]
- Asur, Sitaram, and Bernardo A. Huberman. 2010. Predicting the future with Social Media. Paper presented at 2010 IEEE/WIC/ACM International Conference of Web Intelligence and Intelligent Agent Technology, Toronto, ON, Canada, August 31–September 3.
- Batanero, José María Fernández, Miguel María Reyes Rebollo, and Marta Montenegro Rueda. 2019. Impact of ICT on students with high abilities. Bibliographic review (2008–2018). *Computers and Education* 137: 48–58. [CrossRef]
- Bennett, W. Lance. 2015. Changing citizenship in the digital age. In *Civic Life Online: Learning How Digital Media Can Engage Youth*. Edited by W. Lance Bennett. Cambridge: MIT Press.
- Caballero, Francisco Sierra. 2013. *Ciudadanía, Tecnología y Cultura: Nodos Conceptuales para Pensar la Nueva Mediación Digital*. Barcelona: Gedisa.
- Chen, Hsinchun, Roger HL Chiang, and Veda C. Storey. 2012. Business intelligence and analytics: From Big data to big impact. *MIS Quarterly* 36: 1165–88. [CrossRef]
- Chen, Min, Shiwen Mao, and Yunhao Liu. 2014. Big Data: A Survey. *Mobile Networks and Applications* 19: 171–209. [CrossRef]
- Correa, José-Miguel. 2015. ¿Cómo aprender a ser maestro?: Tic, género y narrativas visuales de futuras maestras de educación infantil. *Reire* 8: 256–68.
- Daniel, Ben Kei. 2015. Big Data and analytics in higher education: Opportunities and challenges. *British journal of Educational Technology* 46: 904–20. [CrossRef]
- Daniel, Ben Kei. 2017. Big Data and data science: A critical review of issues for educational research. *British Journal of Educational Technology* 50: 101–13. [CrossRef]
- DatAnalysis 15M. 2013. Tecnopólitica: La Potencia de las Multitudes Conectadas. El Sistema-Red 15M Como Nuevo Paradigma de la Política Distribuida. Available online: <http://journals.uoc.edu/ojs/index.php/in3-working-paper-series/article/view/1878> (accessed on 29 May 2019).

- Dishon, Gideon. 2017. New data, old tensions: Big data, personalized learning, and the challenges of progressive education. *Theory and Research in Education* 15: 272–89. [\[CrossRef\]](#)
- Dussel, Ines. 2012. *Aprender a Enseñar en la Cultura Digital*. Buenos Aires: Fundación Santillana.
- Dussel, Ines. 2014. ¿Es el currículum relevante en la cultura digital? Debates y desafíos sobre la autoridad cultural contemporánea. *Archivos Analíticos de Política Educativa* 22: 1–22.
- Eichstaedt, Johannes C., Hansen Andrew Schwartz, Margaret L. Kern, Gregory Park, Darwin R. Labarthe, Raina M. Merchant, Sneha Jha, Megha Agrawal, Lukasz A. Dziurzynski, Maarten Sap, and et al. 2015. Psychological language on Twitter predicts county-level heart disease mortality. *Psychological Science* 26: 159–69. [\[CrossRef\]](#)
- Elia, Gianluca, Gianluca Solazzo, Gianluca Lorenzo, and Giuseppina Passiante. 2019. Assessing learners' satisfaction in collaborative online courses through a big data approach. *Computers in Human Behavior* 92: 589–99. [\[CrossRef\]](#)
- Genevieve Bell, Melissa Gregg, and Nick Seaver. 2015. *Data, Now Bigger and Better!* Edited by Tom Boellstorff and Bill Maurer. Chicago: Prickly Paradigm Press.
- Ghani, Norjihhan Abdul, Suraya Hamid, Ibrahim Abaker Targio Hashem, and Ejaz Ahmed. Forthcoming. Social media big data analytics: A survey. *Computers in Human Behavior*. [\[CrossRef\]](#)
- Gibson, David. 2017. Big data in higher education: research methods and analytics supporting the learning journey. *Technology, Knowledge and Learning* 22: 237–41. [\[CrossRef\]](#)
- Glänzel, Wolfgang, and Urs Schoepflin. 1999. A bibliometric study of reference literature in the sciences and social sciences. *Information Processing & Management* 35: 31–44.
- Gorospe, José Miguel Correa, Lorea Fernández Olaskoaga, Aingeru Gutiérrez-Cabello Barragán, Daniel Losada Iglesias, and Begoña Ochoa-Aizpurua Agirre. 2015. Formación del profesorado, tecnología educativa e identidad docente digital. *Revista Latinoamericana de Tecnología Educativa* 14: 46–6.
- Hernández, Dolors Reig. 2012. Socionomia. In *¿Vas a Perder la Revolución Social?* Barcelona: Ediciones Deusto.
- Hicks, Stephanie C., and Rafael A. Irizarry. 2018. A guide to teaching data science. *The American Statistician* 72: 382–91. [\[CrossRef\]](#)
- Hinojo-Lucena, Francisco-Javier, Inmaculada Aznar-Díaz, María-Pilar Cáceres-Reche, and José-María Romero-Rodríguez. 2019. Artificial Intelligence in Higher Education: A Bibliometric Study on its Impact in the Scientific Literature. *Education Science* 9: 51. [\[CrossRef\]](#)
- Huda, Miftachul, Andino Maselena, Masitah Shahrill, Kamarul Azmi Jasmi, Ismail Mustari, and Bushrah Basiron. 2017. Exploring adaptive teaching competencies in big data era. *International Journal of Emerging Technologies in Learning (ijET)* 12: 68–83. [\[CrossRef\]](#)
- Huda, Miftachul, Andino Maselena, Pardimin Atmotiyoso, Maragustam Siregar, Roslee Ahmad, Kamarul Jasmi, and Nasrul Muhamad. 2018. Big data emerging technology: insights into innovative environment for online learning resources. *International Journal of Emerging Technologies in Learning (ijET)* 13: 23–36. [\[CrossRef\]](#)
- Hussain, Amir, and Erik Cambria. 2018. Semi-supervised learning for big social data analysis. *Neurocomputing* 275: 1662–73. [\[CrossRef\]](#)
- Injadat, MohammadNoor, Fadi Salo, and Ali Bou Nassif. 2016. Data mining techniques in social media: A survey. *Neurocomputing* 214: 654–70. [\[CrossRef\]](#)
- Jagadish, Hosagrahar Visvesvaraya. 2016. The values challenge for Big Data. In *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*. Kentucky: IEEE COMPUTER SOCIETY, pp. 77–84.
- Jin, Xiaolong, Benjamin W. Wah, Xueqi Cheng, and Yuanzhuo Wang. 2015. Significance and challenges of big data research. *Big Data Research* 2: 59–64. [\[CrossRef\]](#)
- Khan, Muhammad, Md Karim, and Yangwoo Kim. 2018. A two-stage big data analytics framework with real world applications using spark machine learning and long Short-term memory network. *Symmetry* 10: 485. [\[CrossRef\]](#)
- Liang, Jiajun, Jian Yang, Yongji Wu, Chao Li, and Li Zheng. 2016. Big data application in education: dropout prediction in Edx MOOCs. Paper presented at 2016 IEEE Second International Conference on Multimedia Big Data (BigMM), Taipei, Taiwan, April 20–22.
- Martín-Barbero, Jesús. 2012. Ciudad educativa: De una sociedad con sistema educativo a una sociedad de saberes compartidos. In *Educación Expandida*. Edited by Rubén Díaz and Juan Freire. Sevilla: Zemos98, pp. 103–28.
- Medici, Emilio. 2009. *La Receta de la Industria Creativa Como Motor de Desarrollo y Sus Contradicciones*. *Nuevas Economías de la Cultura*. Madrid: YProductions.

- Menon, Ashwin, Shiv Gaglani, M. Ryan Haynes, and Sean Tackett. 2017. Using “big data” to guide implementation of a web and mobile adaptive learning platform for medical students. *Medical Teacher* 39: 975–80. [CrossRef]
- Merceron, Agathe, Paulo Blikstein, and George Siemens. 2015. Learning analytics: from big data to meaningful data. *Journal of Learning Analytics* 2: 4–8. [CrossRef]
- Moreno-Carriles, Rosa María. 2018. Big data¿ Pero qué es? *Angiología* 70: 191–94. [CrossRef]
- Ni, Lionel M., Haoyu Tan, and Jiang Xiao. 2016. Rethinking big data in a networked world. *Frontiers of Computer Science* 10: 965–67. [CrossRef]
- Oravec, Chesney S., Mustafa Motiwala, Kevin Reed, Tamekia L. Jones, and Paul Klimo Jr. 2019. Big data research in pediatric neurosurgery: Content, statistical output, and bibliometric analysis. *Pediatric Neurosurgery* 54: 85–97. [CrossRef]
- Perlado-Lamo-de-Espinosa, Marta, Natalia Papí-Gálvez, and María Bergaz-Portolés. 2019. Del planificador de medios al experto en medios: El efecto digital de la publicidad. *Comunicar* 27: 105–14. [CrossRef]
- Picciano, Anthony G. 2012. The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks* 16: 9–20. [CrossRef]
- Price, Derek J. 1986. *Little Science, Big Science ... and Beyond*. Nueva York: Columbia University Press.
- Provost, Foster, and Tom Fawcett. 2013. Data science and its relationship to Big Data and data-driven decision Making. *Big Data* 1: 51–59. [CrossRef] [PubMed]
- Pugna, Irina Bogdana, Adriana Duțescu, and Oana Georgiana Stănilă. 2019. Corporate attitudes towards Big Data and its impact on performance management: A qualitative study. *Sustainability* 11: 684. [CrossRef]
- Reidenberg, Joel R., and Florian Schaub. 2018. Achieving big data privacy in education. *Theory and Research in Education* 16: 263–79. [CrossRef]
- Rodríguez-García, Antonio Manuel, Juan Manuel Trujillo, and José Sánchez. 2019. Impact of scientific productivity on digital competence of future teachers: bibliometric approach on Scopus and Web of Science. *Revista Complutense de Educación* 30: 623–46. [CrossRef]
- Saiki, Sachio, Naoki Fukuyasu, Kohei Ichikawa, Tetsuya Kanda, Masahide Nakamura, Shinsuke Matsumoto, Shinichi Yoshida, and Shinji Kusumoto. 2018. A Study of Practical Education Program on AI, Big Data, and Cloud Computing through Development of Automatic Ordering System. Paper presented at 2018 IEEE International Conference on Big Data, Cloud Computing, Data Science & Engineering (BCD), Yonago, Japan, July 10–12.
- Salazar, Javier. 2016. Big Data en la educación. *Revista Digital Universitaria* 1: 1–16.
- Sanchez, Antonio, and Lisa Burnell Ball. 2015. From Big Data to smart data: Teaching data mining and visualization. Paper presented at International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS)), Las Vegas, NV, USA, July 27–30.
- Seufert, Sabine, Christoph Meier, Matthias Soellner, and Roman Rietsche. 2019. A pedagogical perspective on Big Data and learning analytics: A conceptual model for digital learning support. *Technology, Knowledge and Learning*, 1–21. [CrossRef]
- Shum, Simon Buckingham, and Rebecca Ferguson. 2012. Social learning analytics. *Educational Technology and Society* 15: 3–26.
- Sudolska, Agata, Andrzej Lis, and Róża Błaś. 2019. Cloud computing research profiling: Mapping scholarly community and identifying thematic boundaries of the field. *Social Science* 8: 112. [CrossRef]
- Urbizagástegui Alvarado, Rubén. 2016. El crecimiento de la literatura sobre la ley de Bradford. *Investigación Bibliotecológica* 30: 51–72. [CrossRef]
- Veltri, Giuseppe Alessandro. 2017. Big Data is not only about data: The two cultures of modelling. *Big Data & Society* 4: 1–6. [CrossRef]
- Waller, Matthew A., and Stanley E. Fawcett. 2013. Data Science, Predictive analytics, and Big Data: A revolution that will transform supply chain design and management. *Journal of Business Logistics* 34: 77–84. [CrossRef]
- Williams, Matthew L., Pete Burnap, and Luke Sloan. 2017. Crime sensing with Big Data: The affordances and limitations of using open-source communications to estimate crime patterns. *The British Journal of Criminology* 57: 320–40. [CrossRef]

Williamson, Ben. 2015. Governing software: Networks, databases and algorithmic power in the digital governance of public education. *Learning, Media and Technology* 40: 83–105. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).