Hakemli - Süreli - Bilimsel Dergi
İNTERNET UYGULAMALARI ve YÖNETİMİ
Journal of
INTERNET APPLICATIONS and MANAGEMENT
2020 / 11(2)

Big Data Analytics in Higher Education: A Systematic Review *

Yükseköğretim'de büyük veri analitiği: sistematik bir literatür taraması

Zeynep AYTAÇ [1], zduman@aksaray.edu.tr
Hasan Şakir BİLGE [2], bilge@gazi.edu.tr

Received: 07.02.2020; **Accepted**: 16.09.2020

The learning techniques and environment in education has been changing and developing in the last few years. The data obtained from the activities in online learning environments constitute an important data source for improvement and development studies in higher education using big data technologies. This study is based on a review of the literature that focused on the evolving area of big data analytics in higher education. Four groups of stakeholders, namely students, educators, administrators and course developers, in higher education are discussed in this study by utilizing big data and the conceptual model of educational big data analytics. We also discussed different learning environments in a framework in this research. The main objective of this study is to systematically review 40 published articles on big data analytics in higher education in order to determine the subjects of big data analytics in higher education. Based on the findings of the literature review, especially the curriculum development studies were examined and critical findings were discussed.

Keywords: *Big Data, Higher Education, Learning Environments, Curriculum Development*

Eğitimde öğrenme teknikleri ve ortamları son bir kaç yıl içerisinde farklılaşmakta ve gelişmektedir. Çevrimiçi öğrenme ortamlarındaki etkinliklerden elde edilen veriler, büyük veri teknolojileri kullanılarak yükseköğrenimdeki iyileştirme ve geliştirme çalışmaları için önemli veri kaynakları oluşturmaktadır. Bu çalışma gelişmekte olan büyük veri analitiği alanının, yükseköğrenimde literatürün gözden geçirilmesine dayanmaktadır. Bu çalışmada, yükseköğretimde öğrenciler, eğitimciler, yöneticiler ve ders geliştiriciler olmak üzere dört paydaş grubu, büyük veri ve eğitimsel büyük veri analitiğinin kavramsal modeli kullanılarak tartışılmıştır. Ayrıca bu araştırmada farklı öğrenme ortamları da bir yapı içerisinde tartışılmıştır. Bu çalışmanın temel amacı, yüksek öğrenimde büyük veri analitiğinin hangi konulara daha çok yöneldiğini belirlemek için yüksek öğrenimde büyük veri analitiği ile ilgili yayınlanmış 40 makaleyi sistematik olarak incelemektir. Literatür taramasından elde edilen bulgulara dayanılarak, özellikle müfredat geliştirme çalışmaları incelenmiş ve kritik bulgular tartışılmıştır.

Anahtar Kelimeler: *Büyük Veri, Yükseköğretim, Öğrenme Ortamları, Müfredat Geliştirme*

* 29 Mart 2019'da International Instructional Technologies in Engineering Education Conference (IITEEC)'da bildiri olarak sunulmuştur.

[1] Aksaray Üniversitesi, İktisadi İdari Bilimler Fakültesi
[2] Gazi Üniversitesi, Mühendislik Fakültesi

1. INTRODUCTION

Many institutions are operating in a competitive environment because of great number of students and diverse student profiles, national and global economic, social and political changes (Daniel, 2015), technological developments and Internet usage. In addition to traditional learning methods, blended and online learning methods have also become a challenge for many education providers worldwide. Managing the process of learning and teaching through information and communication technology (ICT), including online learning resources (OLR) has become standard practice in many institutions (Anshari, Alas, Yunus, Sabtu, & Hamid, 2016). Online resources address to a large spectrum of information available on the Internet including texts, videos, images, journals, case studies, databases and curriculums (Smith & Rose, 2002). Learning analytics can improve learning and teaching practices by having a good comprehension into the learning process of students to define learning strategies, and as well enabling personal learning activities (Macfadyen, Dawson, Pardo, & Gašević, 2014). In higher education, the users, namely students, teachers, administrators and developers/researchers, utilize the success of students, teachers, and institutions by collecting and analyzing student, departmental and research data with big data and data mining tools.

In addition to the technological developments regarding learning environments (online or face-to-face), departments in higher education institutions have to prepare individuals in a broader sense according to the needs of the sector they will work in. Big data analytics and related technologies have been developing very rapidly in the last decade. However, as technology, tool and applications continue to progress, it is often found that the static nature of the formal curriculum does not meet the requirements. In this study, articles conducted on big data analytics in higher education were reviewed in order to determine the subjects of big data analytics in higher education and examine curriculum developments. This paper is divided into five main parts. In the following section, namely section 2, the basics and characteristics of big data are explained. In section 3, the conceptual model of educational big data analytics, student success with respect to big data, and applications in learning and educational environments are detailed. In section 4, the methodology of the present study and the selection criteria of the articles are given. Finally, in the last section, the findings regarding big data in higher education and the key points of the articles related to curriculum development are discussed.

2. BIG DATA

Many organizations use different kinds of data analysis solutions for decision-making to achieve their operational and strategic goals. How data is stored and analyzed has changed because the nature of the data available in the organizations has changed. It is challenging to manage the new type of data that is unstructured and complex.

The new types of data that are growing very fast and from different sources are not sufficient with traditional database systems. It is evident that the exponential growth of data has brought about tremendous changes in data access, storage and processing in business and therefore, database systems. Big data analytics is an analytical method that can turn unstructured data into useful patterns and information.

Manyika et al. (2011) defined big data as *data that is too big and moves too fast, thus exceeding the processing capacity of conventional database systems*. Storing, managing and analyzing large data sets with different structures requires innovative technologies and techniques (Daniel, 2015). Davis and Patterson (2012) referred to big data as data that is too large to be analyzed by conventional database protocols such as structured query language (SQL). Douglas (2001) determined the three most common properties of big data as volume, velocity and variety in Gartner's report. Given below are the fundamental characteristics of big data (Singh & Singh, 2017; Daniel B., 2015; Zikopoulos, Eaton, deRoos, Deutsch, & Lapis, 2012; Lohr, 2012):

- Volume – is the size of the data consisting of terabytes or petabytes. A large amount of data is difficult to store, transfer and analyze.
- Velocity – is the speed of the data which is generated. Decision-makers want the information to provide data in a fraction of a second in real-time. (e.g., status updates, likes and shares on social media)
- Variety – refers to different data formats, data semantics and data structures like audio, images, sensor data and weblogs.
- Veracity – means the accuracy or truth of the data.
- Value – refers to the ability to analyze data and provide a better understanding of the various key areas.
- Verification – refers to data verification and security.

3. BIG DATA IN EDUCATION

Managing, tracking and controlling learners, courses, grades, degrees, certificates, institutions and learning providers requires a content management system. Educational environments involve different groups of users and these users have knowledge from a wide range of perspectives with respect to their objectives, vision and mission (Margo, 2004).

Romero and Ventura (2010) suggested five groups of users for educational environments: students, educators, course developers, organizations/learning providers and administrators. Siemens and Long (2011) reported two types of analytics: (1) *Learning analytics,* which include course-level and departmental level and (2) *academic analytics,* which include institutional, regional, national, and international level. Higher education can utilize big data analytics for students, educators, administrators and course developers.

Institutions can benefit from big data (Figure 1):

- For students, by personalizing learning, planning learning activities, recommending courses, activities and resources.
- For educators, by analyzing student behavior, predicting student performance and grouping and modeling students.
- For administrators, by organizing resources, enhancing educational programs and evaluating teachers and curricula.
- For course developers, by evaluating courseware, improving student learning, constructing student models and tutor models.

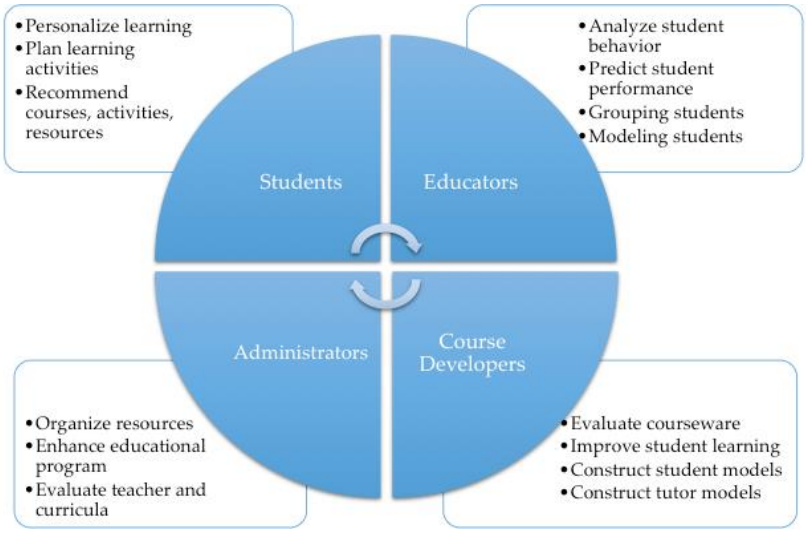

Figure 1. Big data scopes for four users in higher education (Romero & Ventura, 2010)

Daniel and Butson (2013) aimed to find out the guidelines of Big Data within universities and developing frameworks for data aggregation, collection and develop a data set that relates with end-users. The end users are students, instructors, policy and researchers. They suggested a conceptual framework for identifying big data analytics in higher education (Figure 2). The framework consists of four components, namely learning analytics, information technology analytics, academic analytics, institutional analytics. The definitions of the four components are as given below:

- Learning Analytics: is collecting, measuring, analyzing and reporting the data regarding students and their circumstances, with the intend of understanding and improving their learning and the region in which it results (Siemens & Long, 2011).
- Institutional Analytics: analyses operational data to support effective decision making at an institutional level like instructional analytics, assessment policy analytics and structural analytics.
- Information Technology Analytics: uses performance data, which supports the examination of technology, tools, processes, progressing data standards, policies and organizational synergies.
- Academic/Program Analytics: includes information regarding programs and performance challenges.

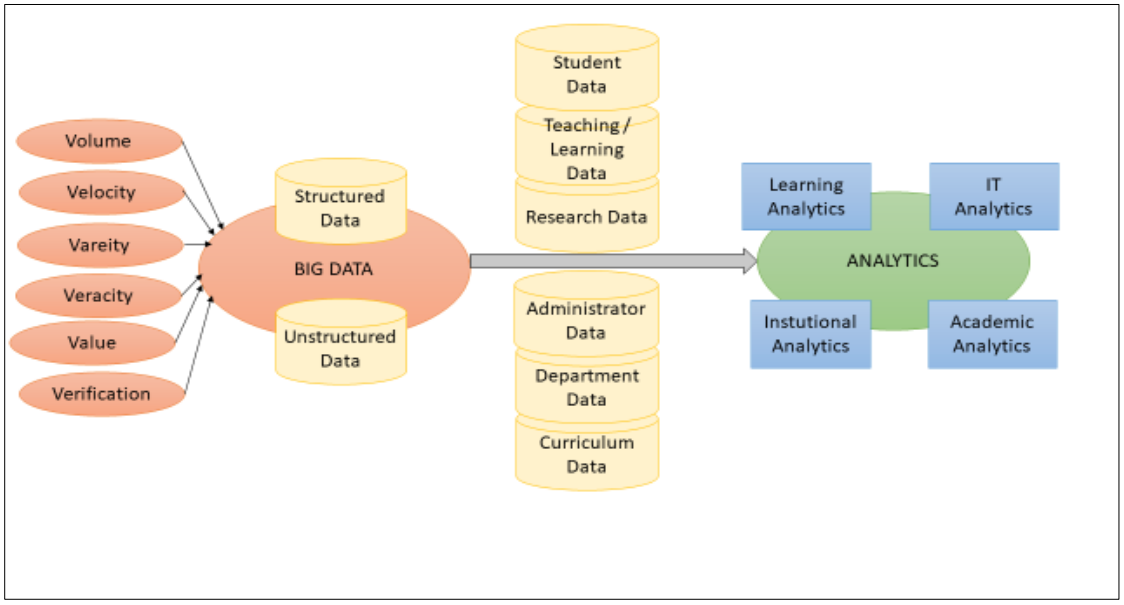

Figure 2. Conceptual model of educational big data analytics adapted from Daniel & Butson (2013)

3.1. Big Data and Student Success

Today, student's activity on cell phones, tablets, search engines, social networks and GPS navigators creates a tremendous amount of information, often in real-time. When and where students shop, eat and engage in activities and lectures can be tracked with the use of student identification cards. Course management software records what materials student access and how long they engage with them. Big data creates knowledge for higher education including the potential to develop customized learning experiences (Lane & Finsel, 2014).

Student success is a result of a complex interacting of many factors such as curriculum, background, funding, testing style, class size and school size. IBM defined some categories to improve student performance by using data in higher education in a white paper. These categories are given below (IBM Software Group, 2001):

1. Measuring and monitoring student achievements
2. Reporting student achievements by levels such as economic status, gender and language proficiency
3. Identifying anomalies for early intervention
4. Specifying potential by revealing patterns for individual student performance
5. Preventing abandonment by building student community and strengthening relationships between teachers and students
6. Defining and improving key attributes of good instructors
7. Examining standardized testing to balance performance
8. Testing and evaluation of curricula

3.2. Applications in Learning

Collecting and analyzing data has changed learning and teaching strategies. Information on what students do, such as whether they attend lectures and how they perform, is important to evaluate their particular strengths and weaknesses.

A lot of methods have been used for educational environments that have been analyzed through data mining and big data. Baker and Yacef (2009) suggested four significant methods, namely improvement of student patterns, working on the pedagogical support, discovering domain models and scientific research into learning, for educational data mining.

Romero and Ventura (2010) determined the following categories of application for main educational tasks:

- Analysis and visualization of data: Instructors and course developers can obtain overall information about students and resolve the course activities of students to understand student's learning. The purpose of the analysis and visualization of data is to support decision making and create useful information.
- Feedback for supporting educators: The purpose of providing feedback to assist teachers, contributers and management team on giving a decision is to develop learners' insight and arrange educational resources more productively.
- Guidance for students: The purpose of guidance is to make recommendations to student regarding new courses by analyzing their personalized activities.
- Predicting student's performance: The purpose of predicting a student's performance is to estimate the undetermined value of a variable that identifies the student. Student's interaction with other instructors and learners in a learning environment can be analyzed and predicted their performance.
- Student modeling: The purpose of student modeling is to create models, including learners' knowledge and skills, with characteristics like satisfaction, motivation, affective status and learning styles.
- Pointing undesirable student behavior: The purpose of pointing undesirable student behavior is to determine the students who show extraordinary attitude like low motivation, erroneous actions, cheating, playing games, academic failure and dropping out.
- Grouping students: The purpose of grouping students is to create groups/clusters by customizing their features and personal characteristics. Instructors build personalized learning systems, promote active group learning and ensure applicable content by grouping the students.
- Social network analysis: The purpose of social network analysis is to analyze the relationships between individuals. A group of people or an organization who are related by social in common such as friendship, collaboration or informational sharing is described as a social network (Freeman, 2004). In educational environments, there are different methods for social network analysis. Collaborative filtering is the most common method for mining social networks.

- Developing concept maps: It is a conceptual graph that demonstrates relationships between concepts and states the hierarchal structure of knowledge (Novak & Canas, 2008). Association rule mining and text mining have been used to build concept maps.
- Constructing courseware: The purpose of constructing courseware is to support educators in the developing process of courseware and learning contents.
- Planning and scheduling: The purpose of planning and scheduling is to improve the conventional educational process by suggesting students with planning, course scheduling, future courses, developing curriculums, helping in the admission processes and planning resource allocation.

3.3. Learning Environments

Technology has been applied to learning for many years and learning environments have evolved quickly with online and mobile learning, learning management systems, tools and technological developments. Bersin (2018) put forward the term 'digital learning' in his article titled 'The Disruption of Digital Learning: Ten Things We Have Learned', which refers to "*delivering learning to where people are*". Digital learning is identified as a '*way of learning'* not a '*type of learning.'* According to Bersin, from traditional learning to e-learning, blended learning, talent-driven learning and continuous learning have evolved extremely fast. In digital learning, traditional learning management systems are no longer the center of learning. Micro-learning platforms support learners at any time and enable them to reach course contents everywhere. The next step is intelligent learning, which is personalized, intelligent and machine-driven (Bersin, 2018). Artificial intelligence-based learning will make a significant contribution to education on activities such as student needs, grading, course improvements, automated tutors and feedback in future learning environments.

There are a variety of educational learning environments like traditional classrooms, online learning, mobile learning, blended learning, digital game-based learning and gesture-based learning (Figure 3). In addition, there are a wide variety of educational systems like learning management systems, learning object repositories, social networks, forums and virtual environments. Processing all data provided by different educational systems and environments can be resolved by using big data analytics and data mining techniques.

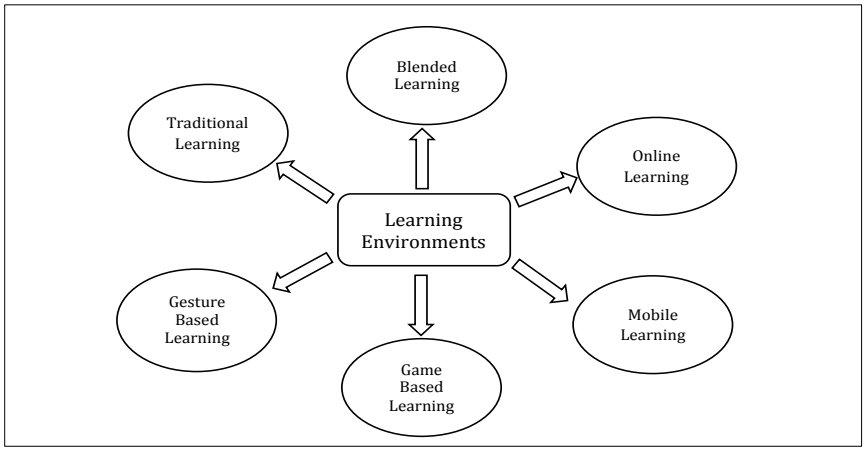

Figure 3. Types of learning environment

3.3.1. Face to Face (Traditional) Learning

In classroom environments, students interact with the instructor through face-to-face contact. Data collected in classroom environments regarding student's performance, curriculum (Romero & Ventura, 2010), attendance, grades, instant feedback can be used as a raw data. In addition, the profile of mood states can be collected in the classroom and analyzed with various statistical techniques. Data gathered in the classroom environment is not as large as the data gathered in online learning environments. Thus, it is not very sufficient to use big data analytics for data that has been created in traditional learning environments.

3.3.2. Blended Learning

Blended learning is composed of face-to-face interaction and web-based technology. Blended learning environments do not have only one source for gathering information, thus it is difficult to implement learning analytics for blended learning environments. Blended learning combines different types of technologies integrated with face-to-face classroom activities (Picciano, Big Data and Learning Analytics in Blended Learning Environments: Benefits and Concerns, 2014).

Picciano (2014) divided the concept of blended learning into four groups and emphasized that they should not be considered as a limiting and absolute declaration.

- Minimal Technology-Face to Face Classroom: Students meet face-to-face. The instructor uses basic technology like web for e-lectures or e-mail.
- Technology Infused-Face to Face Classroom: Students meet face-to-face. The instructor uses technology like tutorials, digital videos and simulations.
- Minimal Technology-Online Environment: Students meet online. The instructor uses basic technology, like electronic bulletin boards.
- Technology Infused- Online Environment: Students meet online. The instructor uses multiple synchronous and asynchronous technologies like blogs, wikis and interactive video conferencing.

3.3.3. Online Learning

E-learning is a term that refers to technology-supported learning rather than other terms like online learning, web-based learning, and technology-mediated learning and distributed learning (McGill & Klobas, 2009). Online learning changed the way of teaching and learning that emerged in the 1990s and early 2000s. Institutions now offer their academic programs online, and millions of students are enrolled in online programs (Picciano, Big Data and Learning Analytics in Blended Learning Environments: Benefits and Concerns, 2014).

New methods for collecting and processing data on course and student activities have been developed for online learning environments. Institutions have access to large data sets to analyze the data, and there is an exponentially increasing amount of online student data to analyze (Swain, 2013). With online courses, data is gathered for each student transaction. Students use course assessment systems, blogs, discussion boards, and thus generating thousands of transactions per course. Analytics is used for this data, which is collected in real-time and analyzed to suggest courses of action.

3.3.4. Mobile Learning

M-learning is the mobile and communication technology used for teaching and learning in educational institutions that both instructors and students can benefit from (Shorfuzzamana, Hossain, Nazir, Muhammad, & Alamri, 2019). Mobile learning enables students to make use of their mobile devices to access learning materials like lecture notes and slides and informative videos from anywhere and at any time. Mobile learning also gives educators more mobility and flexibility for the creation of field learning.

According to Statista's January 2018 data, the globally mobile population was 3.7 billion unique users. In 2019, 63.4% of the global mobile phone users accessed the Internet from their mobile devices (Figure 4). In 2021, worldwide mobile data traffic is expected to reach 47% of the growth rate. The increasing number of mobile phone users and those accessing the Internet from mobile devices shows that mobile learning will become crucial in higher education and attract more and more users. One of the advantages of mobile learning is reducing operating costs for institutions as students bring their own devices.

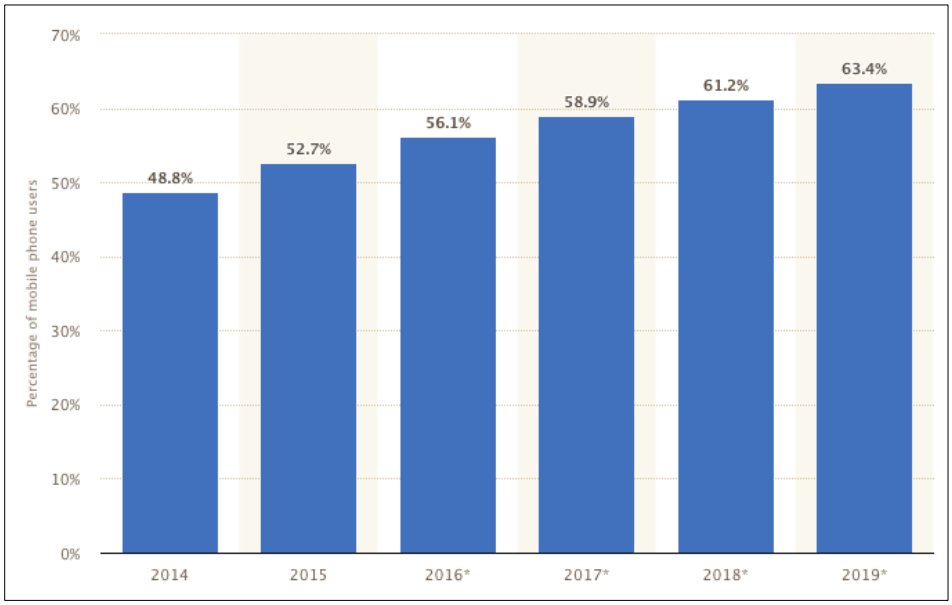

Figure 4. Mobile phone Internet user penetration worldwide from 2014 to 2019 (Statista.com)

3.3.5. Digital Game-Based Learning

The usage of the Internet and communication technology itself are not enough to motivate students. New learning types that correspond better with their interests and habits should be considered to attract students' attention. Within this context, computer games and digital games may be the answer (Prensky, 2003).

Eck (2016) identified three factors of widespread public interest in games as a learning perspective. The first factor is that digital game-based learning is more popular day by day, and researchers tend to conduct game-based learning. The second factor is that the new generation is disconnected from traditional instructions. They have visual skills that are compatible well with digital games, want quick and frequent interactions with content and

need multiple streams of information. The third factor is that the popularity of digital games is increasing (Eck, 2006).

3.3.6. Gesture-Based Learning

Users interact with computers and tools more directly through the use of movements and motions like body movements, facial expressions, finger flips, eye movements, gestures and speech with gesture-based devices (Johnson, Adams, & Cummins, 2012).

The Horizon report has indicated that there are new opportunities in learning by using gesture-based devices as an emerging technology. Innovative teaching and learning with gesture-based systems are increasingly used in the fields of special education, mathematics, physics, physical therapy, music, arts, science, social development and literacy (Johnson, Levine, Smith, & Stone, 2010).

4. METHOD

In this systematic literature review, articles conducted on the subjects of big data in higher education and curriculum development in higher education were selected. In the literature review, the following criteria were used to select the articles:

- Being related to big data in higher education.
- Having used the Scopus database for research.
- Having been published between 2010 and 2020.
- Having been published in English.

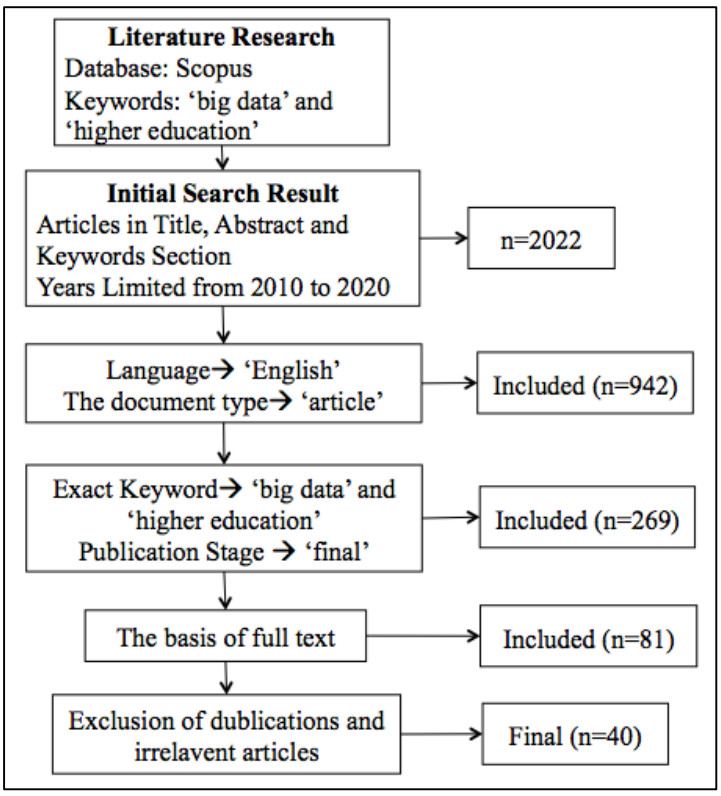

Figure 5. Article Selection Process

The following steps were used for the selection of the research material (Figure 5):

1. The keywords 'big data' and 'higher education' were used. As a result, a total of 269 relevant articles were found. Only open access articles were selected, thus, the number of articles decreased to 62. Non-related articles were also eliminated. In the end, 40 articles were examined within the scope of this study.
2. The search was conducted in the title, abstract and keywords field of the Scopus database.
3. The document type was selected as article. Dissertation and book chapters were excluded, while articles in the final stage of publication were included
4. The keywords were sorted by relevance with the Scopus Advanced Search.

5. FINDINGS

The article selection process and the number of articles included in each step according to the criterion are shown in Figure 5. After the articles that included the keywords "big data" and "higher education" and those that were in the final publication stage were selected a total of 269 articles were determined. These articles were then screened to select those that were full text and as a result 81 articles were obtained. Then, these articles were examined in terms of title, abstract and keywords. The articles discussing big data analytics with regards to any sector other than higher education were excluded. Those discussing higher education within the scope of different approaches than big data analytics were also excluded. As a result, a total of 40 articles were included in this review.

The literature included in this review mainly covered teaching and learning analytics, distance learning and the challenges and opportunities of big data in higher education. Other articles were related to curriculum development, ethics and privacy, assessment, adoption, visualization, education policy and predictive models. In total 12 of the reviewed articles were regarding learning and teaching analytics. It was determined that in these articles the most used methodologies were literature review or framework development and model designing. In addition, case study and thematic analysis had been used as the methodology of these articles. A total of six articles were related to distance learning. Distance learning was also referred to as online learning or e-learning in the articles. Case study, review, model designing, framework development, clustering algorithm and chaos optimization methods had been used in the articles related to distance learning. The third most mentioned subject regarding big data in higher education was challenges and opportunities. All of the reviewed articles and their methods, subject areas, titles, authors and publishing years are shown in Table1.

Table 1. Big data in Higher Education Subject Areas and Methods

Subject Areas	Number of Articles	Method	Article Titles	Author-Year
Challenges / Opportunities	6	Literature Review	1. Uncertainty in big data analytics: survey, opportunities, and challenges	(Hariri, Fredericks, & Bowers, 2019)
			2. Big Data and analytics in higher education: Opportunities and challenges	
			3. The Dynamic Landscape of Higher Education: The Role of Big Data and Analytics	(Mahroeian & Daniel, 2016)
			4. Impacts on Statistics Education in Big Data Era	(Li, Huang, & Zhou, 2018)
			5. Opportunities and challenges for big data analytics in US higher education: A conceptual model for implementation	(Attaran, Stark, & Stotler, 2018)
			6. Big Data Management, Data Science and Data Analytics: What is it and Where— An Educational in Indian Perspective	(Gokul, Sundararajan, & Paul, 2019)
Distance Learning (Online Learning / E-Learning)	6	1. Case Study	1. Big(ger) Data as Better Data in Open Distance Learning	(Prinsloo, Archer, Barnes, Chetty, & Zyl, 2015)
		2. Clustering Algorithm	2. Research on Cloud Computing and Its Application in Big Data Processing of Distance Higher Education	(Zhang, Li, & Hao, 2015)
		3.Framework Development	3. UTiLearn: A Personalised Ubiquitous Teaching and Learning System for Smart Societies	(Mehmood, Alam, Albogami, Katib, Albeshri, & Altowaijri, 2017)
		4. Model Designing	4. Big Data Emerging Technology: Insights into Innovative Environment for Online Learning Resources	(Huda, et al., 2018)
		5. Literature Review	5. Understand, develop and enhance the learning process with big data	(Sedkaoui & Khelfaoui, 2019)
		6. Chaos Optimization	6. A cognitive learning model in distance education of higher education institutions based on chaos optimization in big data environment	(Wen, Zhang, & Shu, 2019)
Curriculum Development	4	1. Qualitative & Quantitative	1. Business Intelligence and Big Data in Higher Education: Status of a Multi-Year Model Curriculum Development Effort for Business School Undergraduates, MS Graduates, and MBAs	(Gupta, Goul, & Dinter, 2015)
		2. Case Study	2. Sustaining Employability: A Process for Introducing Clod Computing, Big Data, Social Networks, Mobile Programming and Cybersecurity into Academic Curricula	(Bologa, Lupu, Boja, & Georgescu, 2017)
		3. Qualitative (Survey)	3. Measuring the promise of Big Data syllabi	(Friedman, 2018)
		4.Framework Development	4. Data Analytics Research-Informed Teaching in a Digital Technologies Curriculum	(Lu, 2020)
Teaching and Learning Analytics	12	1. Review	1. The Evolution of Big Data and Learning Analytics in American Higher Education	(Picciano, 2012)
		2. Case Study	2. Application of Learning Analytics in University Mathematics Education	(Kim, Hue, & Shin, 2016)
		3. Thematic Analysis	3. Innovative Teaching In Higher Education: The Big Data Approach	(Huda, et al., 2016)
		4. Review	4. Inferring Learning from Big Data: The Importance of a Transdisciplinary and Multidimensional Approach	(Lodge, Alhadad, Lewis, & Gasevic, 2017)

		Table 1. (Continued)		
		5. Conceptual Research Design	5. Impacting Big Data analytics in higher education through Six Sigma techniques	(Laux, Li, Seliger, & Springer, 2017)
		6. Model Designing	6. Understanding Modern Learning Environment (MLE) in Big Data Era	(Huda, et al., 2018)
		7.Framework Development	7. A Framework for Managing and Analyzing Big Data in Indian School Education System with Reference to Jammu & Kashmir	(Kamal & Dave, 2019)
		8. Review	8. A Comparative Study on Big Data Applications in Higher Education	(Altaye & Nixon, 2019)
		9. Infrastructure Design	9. A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education	(Moscoso-Zea, Castro, Paredes-Gualtor, & Lujan-Mora, 2019)
		10. Systematic Literature Review	10. Exploring Data Driven Youth Character Education Frameworks: A Systematic Literature Review on Learning Analytics Models and Participatory Design	(Baiocchi, 2019)
		11. Framework Development	11. Personalized Learning Environment in Higher Education through Big Data and Blended Learning Analytics	(Tulasi & Suchithra, 2019)
		12. Review	12. The Digitalisation of the Economy and Higher Education	(Klochkova, Serkina, Prasolov, & Movchun, 2020)
Ethics / Privacy	3	1. Review/ Discussion	1. Ethical Considerations in the Practical Application of the Unisa Socio-Critical Model of Student Success	(Fynn, 2016)
		2. Review	2. Achieving big data privacy in education	(Reidenberg & Schaub, 2018)
		3. Model Designing	3. Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy	(Jones, 2019)
Education Policy and Challenges	1	Review	Education Policy Research in the Big Data Era: Methodological Frontiers, Misconceptions, and Challenges	(Wang, 2017)
Visualization	2	1.Weighted Multi-dimensional Scaling	1. Be the Data: Embodied Visual Analytics	(Chen, et al., 2017)
		2. Mixed Method (Survey, Self-assessment)	2. Visualization and Experiential Learning of Mathematics for Data Analytics	(Venkatraman, Overmars, & Wahr, 2019)
Adoption	2	1. Model Designing	1. Big Data Analytics Model for the Education Sector	(Shah & Choksi, 2019)
		2. Review	2. Emerging Trends and Future Perspective of Human Resource Reskilling in Higher Education	(Suhasini & Kumar N , 2019)
Assessment	3	1.Framework Development with R, BigQuery	1. Big Data Technology-Enabled Analytical Solution for Quality Assessment of Higher Education Systems	(Khan, Liu, Shakil, & Alam, 2019)
		2. Qualitative Case Study	2. Assessment big data in Nigeria: Identification, generation and processing in the opinion of the experts	(Esomonu, Esomonu, & Eleje, 2020)
		3. Qualitative Case Study	3. Value Creation of Big Data Utilization: The Next Frontier for Productive Scholarship among Filipino Academics	(Dapiton & Canlas, 2020)
Predictive Models	1	Qualitative Case Study	"We called that a behavior": The making of institutional data	(Whitman, 2020)

Four articles carried out on curriculum development were also reviewed. One of them provided a new curriculum model for Business Information Systems undergraduate, MS and MBA students. Gupta, Goul and Dinter (2015) brought together faculty, students and industry experts to add appropriate elective courses to the existing curriculum using Krathwohl's revised taxonomy. They used a multi-methodological approach through interviews, surveys, focus groups and literature reviews. Krathwol's revised taxonomy uses two dimensions, namely knowledge and cognitive process, to the identify learning outcomes in a curriculum. They evaluated Krathwohl's two dimensions on the BI topics in undergraduate, MS and MBA curricula. They developed BI topics across three levels of BI model curricula. Bologa et al. (2017) introduced an academic curriculum including topics such as cloud computing, big data, social networks, mobile programming and cyber security for non-technical students. Their study was funded by the European Union and involved 61 experts including professors, students, company representatives and technical staff members. After their pilot test, two partner universities updated their undergraduate and graduate programs. Their study found that non-technical students studying business, finance, economics, statistics, marketing and similar disciplines provided good feedback and showed interests in the new curricula. Universities could add these topics to the curricula of their undergraduate and graduate programs.

Friedman (2018) examined the content of Big Data syllabi and the major topics covered in these syllabi across different academic institutions in the USA. A total of 35 syllabi were collected from both private and public universities and most were offered in the field of Information Technologies, namely Computer Science, Engineering and Information Science. In addition, the field of Business offered courses on big data too. The eight most popular topics were determined as Big Data Infrastructure (%56), Data Driven Application Systems (%49), Hadoop MapReduce and Spark, R and Python, Data Mining Models (%41), Data Sources, Machine Learning, Statistical Analytics, Predictive Analytics and Visualization. As a result of the study it was determined that there was a lack of standardization with regards to the textbooks, resources and core subjects of big data. Lu (2020) introduced a data-driven analytical framework that provided the architecture for the development of a data analytics curriculum. In the study, a data analytics pathway to evaluate the curriculum from a teaching, learning and assessment perspective was put forward. The area of curriculum development on big data analytics has mostly been aimed at non-technical students, especially business students in higher education. However, it has been made clear that big data analytics are needed in many departments. Analytical abilities and skills should be clearly identified according to the expectations of the industry and big data analytics should be included in the curriculum of different departments of universities.

6. CONCLUSION

Big data analytics plays an important role in higher education. In this study, big data and its characteristics were summarized. Four stakeholders, namely students, educators, administrators and course developers, and their gains were discussed by using big data technologies. Planning learning activities, recommending courses, analyzing student behavior, predicting student performance, grouping and modeling students, enhancing educational programs, evaluating teachers, curricula and courseware, improving student

learning and constructing student models and tutor models are the critical benefits for the four stakeholders. The conceptual model of educational big data analytics was also reviewed and discussed. This study also examined how big data technologies contribute to learning analytics and student success. The different types of learning environments such as traditional classrooms, online learning, mobile learning, blended learning, digital game-based learning and gesture-based learning were also briefly mentioned in a framework.

Big data analytics and related technologies have been developing rapidly in the last decade. However, formal curricula do not meet the industry requirements. In this study, 40 articles on big data analytics in higher education were systematically reviewed in order to determine the areas of big data analytics in higher education and examine curriculum development studies. Based on the findings of the literature review, the most published articles on big data analytics in higher education were in the field of learning and teaching analytics. Literature review and framework development methods had been used in most of these articles. The second and third most published articles on big data in higher education were in the field of online learning and the challenges and opportunities of big data in higher education. In addition, it should be noted that there were not many published articles on curriculum development. Based on the findings of the literature review, it can be said that the area of curriculum development on big data analytics has mostly been aimed at non-technical students in higher education. However, it has been determined that there is a need for big data analytics in many degrees as well as information systems.

REFERENCES

Altaye, A. A., & Nixon, J. S. (2019). A Comparative Study on Big Data Applications in Higher Education. *International Journal of Emerging Trends in Engineering Research*, *7*(12), 739-745.

Anshari, M., Alas, Y., Sabtu, N. P. H., & Hamid, M. S. A. (2016). Online Learning: Trends, Issues and Challenges in the Big Data Era. *Journal of e-Learning and Knowledge Society*, *12*(1), 121-134.

Attaran, M., Stark, J., & Stotler, D. (2018). Opportunities and Challenges for Big Data Analytics in Us Higher Education: A Conceptual Model for Implementation. *Industry and Higher Education*, *32*(3), 169-182.

Baiocchi, R. R. (2019). Exploring Data Driven Youth Character Education Frameworks: A Systematic Literature Review on Learning Analytics Models and Participatory Design. *Estudios Sobre Educacion*, *37*, 179-198.

Baker, R. S., & Yacef, K. (2009). The State of Educational Data Mining In 2009: A Review and Future Visions. *JEDM| Journal of Educational Data Mining*, *1*(1), 3-17.

Bersin, J. (2018, 03 31). https://joshbersin.com. Retrieved 11 26, 2018, from https://joshbersin.com: https://joshbersin.com/2017/03/the-disruption-of-digital-learning-ten-things-we-have-learned/

Bologa, R., Lupu, A. R., Boja, C., & Georgescu, T. M. (2017). Sustaining Employability: A Process for Introducing Cloud Computing, Big Data, Social Networks, Mobile Programming and Cybersecurity into Academic Curricula. *Sustainability*, *9*(12), 2235.

Chen, X., Self, J. Z., House, L., Wenskovitch, J., Sun, M., Wycoff, N., & North, C. (2017). Be the Data: Embodied Visual Analytics. *IEEE Transactions on Learning Technologies*, *11*(1), 81-95.

Daniel, B. (2015). Big Data and Analytics in Higher Education: Opportunities and Challenges. *British Journal of Educational Technology*, *46*(5), 904-920.

Daniel, B. K., & Butson, R. (2013). *Technology Enhanced Analytics (TEA) in Higher Education*. International Association for the Development of the Information Society, 89-96.

Dapiton, E. P., & Canlas, R. B. (2020). Value Creation of Big Data Utilization: The Next Frontier for Productive Scholarship among Filipino Academics. *European Journal of Educational Research*, *9*(1), 423-431.

Esomonu, N. P. M., Esomonu, M. N., & Eleje, L. I. (2020). Assessment Big Data in Nigeria: Identification, Generation and Processing in the Opinion of the Experts. *International Journal of Evaluation and Research in Education*, *9*(2), 345-351.

Freeman, L. (2004). The Development of Social Network Analysis. *A Study in the Sociology of Science*, *1*, 687, 159-167.

Friedman, A. (2018). Measuring the Promise of Big Data Syllabi. *Technology, Pedagogy and Education*, *27*(2), 135-148.

Fynn, A. (2016). Ethical Considerations in the Practical Application of the Unisa Socio-Critical Model of Student Success. *The International Review of Research in Open and Distributed Learning*, *17*(6).

Gokul, K., Sundararajan, M., & Paul, P. (2019). Big Data Management, Data Science and Data Analytics: What is it and Where— An Educational in Indian Perspective. *International Journal of Innovative Technology and Exploring Engineering* (IJITEE), 8(12).

Gupta, B., Goul, M., & Dinter, B. (2015). Business Intelligence and Big Data in Higher Education: Status of a Multi-Year Model Curriculum Development Effort for Business School Undergraduates, MS Graduates, and MBAs. *Communications of the Association for Information Systems*, *36*, Article 23.

Hariri, R. H., Fredericks, E. M., & Bowers, K. M. (2019). Uncertainty in Big Data Analytics: Survey, Opportunities, and Challenges. *Journal of Big Data*, *6*(1), 44.

Huda, M., Anshari, M., Almunawar, M. N., Shahrill, M., Tan, A., Jaidin, J. H., & Masri, M. (2016). Innovative Teaching in Higher Education: The Big Data Approach. *TOJET*, November, Special Issue for INTE 2016, 1210-1216.

Huda, M., Maseleno, A., Atmotiyoso, P., Siregar, M., Ahmad, R., Jasmi, K., & Muhamad, N. (2018). Big Data Emerging Technology: Insights into Innovative Environment for Online Learning Resources. *International Journal of Emerging Technologies in Learning (iJET)*, *13*(1), 23-36.

Huda, M., Maseleno, A., Teh, K. S. M., Don, A. G., Basiron, B., Jasmi, K. A., & Ahmad, R. (2018). Understanding Modern Learning Environment (MLE) in Big Data Era. *International Journal of Emerging Technologies in Learning (iJET)*, *13*(05), 71-85.

IBM Software Group. (2001). Analytics for Achievement. Ottawa: IBM Canada.

Johnson, L., Adams, S., & Cummins, M. (2012). The 2012 Horizon Report. Austin: The New Media Consortium.

Johnson, L., Levine, A., Smith, R., & Stone, S. (2010). The 2010 Horizon Report. The New Media Consortium.

Jones, K. M. (2019). Learning Analytics and Higher Education: A Proposed Model for Establishing Informed Consent Mechanisms to Promote Student Privacy and Autonomy. *International Journal of Educational Technology in Higher Education*, *16*(1), 24.

Khan, S., Liu, X., Shakil, K. A., & Alam, M. (2019). Big Data Technology-Enabled Analytical Solution for Quality Assessment of Higher Education Systems. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *10*(6), 292-304.

Kim, D. R., Hue, J. P., & Shin, S. S. (2016). Application of Learning Analytics in University Mathematics Education. *Indian Journal of Science and Technology*, *9*(46), 1-5.

Klochkova, E., Serkina, Y., Prasolov, V., & Movchun, V. (2020). The Digitalisation of the Economy and Higher Education. *Space and Culture, India*, *7*(4), 70-82.

Lane, J. E., & Finsel, B. A. (2014). Fostering Smarter Colleges and Universities. *Building a Smarter University: Big Data, Innovation, and Analytics*, 3-26.

Laux, C., Li, N., Seliger, C., & Springer, J. (2017). Impacting Big Data Analytics in Higher Education through Six Sigma Techniques. *International Journal of Productivity and Performance Management*, *66*(5), 662-679.

Li, Y., Huang, C., & Zhou, L. (2018). Impacts on Statistics Education in Big Data Era. *Educational Sciences: Theory & Practice*, *18*(5), 1236-1245.

Lodge, J. M., Alhadad, S. S., Lewis, M. J., & Gašević, D. (2017). Inferring Learning from Big Data: The Importance of a Transdisciplinary and Multidimensional Approach. *Technology, Knowledge and Learning*, *22*(3), 385-400.

Lohr, S. (2012, 02 11). The Age of Big Data. Big Data's Impact in the World, pp. 1-5.

Lu, J. (2020). Data Analytics Research-Informed Teaching in a Digital Technologies Curriculum. *INFORMS Transactions on Education*, *20*(2), 57-72.

Macfadyen, L. P., Dawson, S., Pardo, A., & Gaševic, D. (2014). Embracing Big Data in Complex Educational Systems: The Learning Analytics Imperative and the Policy Challenge. *Research & Practice in Assessment*, *9*, 17-28.

Mahroeian, H., & Daniel, B. K. (2016, June). The Dynamic Landscape of Higher Education: The Role of Big Data and Analytics. In *EdMedia+ Innovate Learning* (pp. 1320-1325). Association for the Advancement of Computing in Education (AACE).

Margo, H. (2004). Data Mining in the E-Learning Domain. *Campus-Wide Information Systems*, *21*(1), 29-34.

McGill, T. J., & Klobas, J. E. (2009). A Task–Technology Fit View of Learning Management System Impact. *Computers & Education*, *52*(2), 496-508.

Mehmood, R., Alam, F., Albogami, N. N., Katib, I., Albeshri, A., & Altowaijri, S. M. (2017). UTiLearn: A Personalised Ubiquitous Teaching and Learning System for Smart Societies. *IEEE Access*, *5*, 2615-2635.

Moscoso-Zea, O., Castro, J., Paredes-Gualtor, J., & Luján-Mora, S. (2019). A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence & Analytics for Knowledge Management in Education. *IEEE Access*, *7*, 38778-38788.

Novak, J. D., & Cañas, A. J. (2008). The Theory Underlying Concept Maps and How to Construct and Use Them. Technical Report IHMC CmapTools, Florida.

Picciano, A. G. (2012). The Evolution of Big Data and Learning Analytics in American Higher Education. *Journal of asynchronous learning networks*, *16*(3), 9-20.

Picciano, A. G. (2014). Big Data and Learning Analytics in Blended Learning Environments: Benefits and Concerns. *IJIMAI*, *2*(7), 35-43.

Prensky, M. (2003). Digital Game-Based Learning. *Computers in Entertainment (CIE)*, *1*(1), 21-21.

Prinsloo, P., Archer, E., Barnes, G., Chetty, Y., & Van Zyl, D. (2015). Big (Ger) Data as Better Data in Open Distance Learning. *International Review of Research in Open and Distributed Learning*, *16*(1), 284-306.

Reidenberg, J. R., & Schaub, F. (2018). Achieving Big Data Privacy in Education. *Theory and Research in Education*, *16*(3), 263-279.

Romero, C., & Ventura, S. (2010). Educational Data Mining: A Review of the State Of The Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*(6), 601-618.

Sedkaoui, S., & Khelfaoui, M. (2019). Understand, Develop and Enhance the Learning Process with Big Data. *Information Discovery and Delivery*, *47*(1),2-16.

Shah, B., & Choksi, D. (2019). Big Data Analytics Model for the Education Sector. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(12).

Shorfuzzaman, M., Hossain, M. S., Nazir, A., Muhammad, G., & Alamri, A. (2019). Harnessing the Power of Big Data Analytics in the Cloud to Support Learning Analytics in Mobile Learning Environment. *Computers in Human Behavior*, *92*, 578-588.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, *46*(5), 30.

Singh, D. S., & Singh, G. (2017). Big data – A Review. International Research Journal of Engineering and Technology, *4*(4), 822-824.

Smith, A., & Rose, R. (2002). Essential Elements: Prepare, Design, and Teach Your Online Course. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 2723-2724). Association for the Advancement of Computing in Education (AACE).

Suhasini, B., & Kumar N., S. (2019). Emerging Trends and Future Perspective of Human Resource Reskilling in Higher Education. International Journal of Recent Technology and Engineering (IJRTE), 8(2S4).

Swain, H. (2013, August 5). The Guardian. Retrieved October 15, 2018, from http://www.theguardian.com/education/2013/aug/05/electronic-data-trail-huddersfield-loughborough-university

Tulasi, B., & Suchithra, R. (2019). Personalized Learning Environment in Higher Education through Big Data and Blended Learning Analytics. International Journal of Recent Technology and Engineering (IJRTE), 8(3).

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE review*, *41*(2), 16.

Venkatraman, S., Overmars, A., & Wahr, F. (2019). Visualization and Experiential Learning of Mathematics for Data Analytics. *Computation*, *7*(3), 37.

Wang, Y. (2017). Education Policy Research in the Big Data Era: Methodological Frontiers, Misconceptions, and Challenges. *Education Policy Analysis Archives*, *25*(94), 1-24.

Wen, J., Zhang, W., & Shu, W. (2019). A cognitive learning model in distance education of higher education institutions based on chaos optimization in big data environment. *The Journal of Supercomputing*, *75*(2), 719-731.

Whitman, M. (2020). "We Called That a Behavior": The Making of Institutional Data. *Big Data & Society*, *7*(1), 1-13.

Zhang, G., Li, J., & Hao, L. (2015). Research on Cloud Computing and Its Application in Big Data Processing of Distance Higher Education. *International Journal of Emerging Technologies in Learning*, *10*(8), 55-58.

Zikopoulos, P., & Eaton, C. (2011). *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. McGraw-Hill Osborne Media.