Special aspects of using Big Data in the learning process

Olga Agatova, Alexander Popov & Suad Abdalkareem Alwaely

Published online: 19 Sep 2022.

Submit your article to this journal ⬈

View related articles ⬈

View Crossmark data ⬈

Routledge
Taylor & Francis Group

Check for updates

Special aspects of using Big Data in the learning process

Olga Agatova[a], Alexander Popov[b,c,d,e] and Suad Abdalkareem Alwaely[f,g]

[a]Digital Educational Environment Development Laboratory, Russian Academy of Education, Moscow, Russian Federation; [b]Research sector "Open Education", Research Center for Socialization and Resonalization of Children's Education of theFederal Institute for the Development of Education of the Russian Presidential Academy of National Economy and Public Administration, Moscow, Russian Federation; [c]Laboratory for the Development of the Content of Education, the Center for the Development of Education of the Russian Academy of Education, Moscow, Russian Federation; [d]Laboratory of Competence Practices of Education, the Research Institute of Urbanism and Global Education of the Moscow City Pedagogical University, Moscow, Russian Federation; [e]Department of Sociology and Mass Communications, Faculty of Humanities of Novosibirsk State Technical University, Novosibirsk, Russian Federation; [f]Master's Department in Arabic Language Curricula and Islamic Education, Al Ain University, Abu Dhabi, UAE; [g]Department of Curriculum and Instruction, Hashemite University, Zarqa, Jordan

ABSTRACT

The paper examines the special aspects of using Big Data technology in education. The population was made up of 356 third-year university students. To study Big Data technology, a questionnaire was used where respondents rated: cloud technology; apps; Massive Open Online Courses (MOOCs) and digital learning platforms. The study suggested that the education sector is ambitiously applying Big Data technology, both online and offline. All surveyed respondents use apps in Big Data learning and analysis: 73.03% use Moodle, 67.13% use Zoom, 65.17% use Quizlet, 50.84% use Skype, and 35.11% use Slack. MOOCs in education are used by 75% of respondents. Digital learning platforms are used by all respondents. All students use cloud technology. When dealing with Big Data technologies, students preferred apps (8.9 ± 1.33) instead of the cloud (6.9 ± 0.11). Students believe that the important factors for using Big Data in the learning process include: quality of information (85.96%); interest (77.81%); instructor's support (66.85%). The research findings make it possible to integrate Big Data technology into the learning process, thus improving learning outcomes and providing greater speed in processing reliable and meaningful data.

Introduction

The integration of innovative technologies, digitalization and digital transformation of all facets of life and activities led to the emergence of a large amount of information and various data, both offline and online (Baig et al., 2020). Due to the huge amount of information, digital, mobile and communication opportunities, a large number of people today use the Internet and cutting-edge devices in everyday life. According to Statista, the penetration rate of smartphones and other technology in our lives amounts to 65.34%. The highest levels of integration are observed in Saudi Arabia (86.1%), with 5.2 billion Internet connections. Today more than 80% of the population uses the Internet (Shorfuzzaman et al., 2019).

Over the past 30 years, the amount of information and available data has increased exponentially, resulting in a large amount of data (Big Data) and even causing information overload

---

(Saxena & Lamest, 2018). Information overload is a major problem of modern times. The information overload phenomenon has been known by different names: information overabundance, infobesity, infoglut, data smog, information pollution, information fatigue, social media fatigue, social media overload, information anxiety, library anxiety, infostress, infoxication, reading overload, communication overload, cognitive overload, information violence, and information assault (Bawden & Robinson, 2020). Researchers are concerned about the large amount of information that humanity encounters every day (Bawden & Robinson, 2020).

Big Data technology is being actively used in healthcare, insurance, construction, telecommunication, e-commerce (Fiofanova, 2021a). Dresner Advisory Services argues that Big Data are mostly generated by technology (14%), financial services (10%), consulting (9%), healthcare (9%), education (8%) and telecommunication (7%) (Dresner Advisory Services, 2017).

The education sector is the most ambitious creator and consumer of Big Data. Online courses and formats are actively used to support the learning process in today's environment, and especially in the context of the COVID-19 pandemic (Oi et al., 2017). Relying on Big Data, educators are widely expanding the horizons of teaching opportunities, resulting, particularly, in:

- improvements in students' skills and academic achievements;
- feedback in the form of grades, drawing attention to the course content;
- feedback in the form of comments, revealing students' problems and allowing them to improve their knowledge (Black & Wiliam, 2018).

Educators can develop, improve student motivation, and learning outcomes through perceived ease of use and value (Zheng & Bender, 2019). In the current context, education adapts to the students' needs and demands. Many educational websites and online courses are being developed and used to support the learning process (e.g. Quality Matters, MOOCs, community discussion boards, email listservs, Internet browser use, digital learning objects, websites, blogs, Facebook, and others). Such websites and online courses take into account students' preferences (Fiofanova, 2021a; Holland, 2019).

Education is increasingly relying on Big Data and digital opportunities (Shirinkina, 2020). The US has the huge market for online education services, with more than 3 million students (15% of all potential students). HolonIQ predicts that by 2025 more than 1 billion students will have acquired a university degree through digital tools, with this number exceeding 280 million over the next 10 years (Shirinkina, 2020). To effectively process and analyze the massive data sets, Big Data uses cutting-edge technology and apps (Kalaian et al., 2019). Hadoop, MapReduce, NoSQL and other technologies are used to evaluate and personalize Big Data in the learning environment. This educational technology integrates into the learning process such measures as flexibility, scalability, accessibility, security, confidentiality, ease of use, and high quality of content (Mamedova et al., 2017). Especially relevant is the use of Big Data in E-learning environment, which also contributes to the development of students as stakeholders (Moharm & Eltahan, 2020).

Available studies have shown that Big Data in education cover primarily four areas:

(1) learner's performance;
(2) modeling and educational data warehouse;
(3) improvement in the educational system; and
(4) integration of Big Data into the curriculum (Fiofanova, 2021b).

Despite the available studies of Big Data, detailed research of Big Data's current role in education is still lacking. Therefore, a study of special aspects of Big Data technology in education has been conducted.

Literature review

The education sector is ambitiously integrating Big Data technology into the learning process (Ershova et al., 2021). The application of Big Data technology in education has been studied by researchers in China (2021). Currently, educators extensively use big data, which requires the proper infrastructure layer, data resource layer, and service application layer. Using Big Data in learning requires enhancing course content and resources, overhaul of learning models, and use of Big Data analysis. Big Data technology facilitates the integration of M-learning, relevant apps, and provides improved and personalized learning through the provision of educational services tailored to the students' skills and abilities (Qu, 2021).

Educational Big Data was studied by a team of researchers from Taiwan, Japan, Australia, USA (2020). The use of Big Data technology in education is an affordable tool, which makes it possible to:

- assess knowledge automatically;
- compare behavior patterns with personality factors (achievements, learning styles, motivation);
- carry out data mining with new environments (educational games, MOOCs);
- evaluate the academic program's effectiveness;
- ensure confidentiality and security;
- provide support for educators;
- visualize learning activities (Chen et al., 2020).

A team of researchers from the Republic of China, Taiwan and Australia analyzed Educational Big Data from 2010 to 2018 (2020). Big Data technology in education is described by its purpose, problems and methods of analysis. The educational goal: high quality learning. The problems include:

- lack of modeling of student behavior and misuse of resources;
- failure to harmonize academic programs;
- proper quality assurance in education;
- confidentiality and ethical issues.

Analytical methods of Big Data include educational data mining and learning analysis (Quadir et al., 2020).

Big Data technology for higher education has been studied by researchers in Australia and USA (2020). Students prefer the more familiar Big Data tools. When it is possible to use Moodle, they prefer WhatsApp – a more familiar informal e-learning tool. The introduction of Big Data in education will provide greater access to quality education through modern designing and developing instructional resources (Miah et al., 2020).

The use of Big Data and artificial intelligence in education were studied by a team of researchers from Taiwan, Japan, Canada, USA, Brazil, Hong Kong (2020). The flexibility, convenience, and integration of data collection, combined with analytical methods, have allowed education to reach a new level and introduce personalized learning and approach in education (Luan et al., 2020).

Opportunities and challenges of Big Data in education were studied by a team of researchers from Germany, USA, Canada (2020). Big Data and the digital learning environment need newer methods of learning instead of traditional ones. The Big Data environment has a micro level (simple data, such as click data), a meso level (text data) and a macro level (institutional data). Learning management systems, self-regulated learning, and MOOCs contribute to the implementation of Big Data technology. Challenges currently faced by the education sector are attributed to accessing, analyzing, and using Big Data. Big Data problems in education may be solved relying on the benefits of mining Big Data in education:

- forecasting;
- making logical conclusions;
- knowledge modeling;
- establishing connections;
- pattern analysis and correlation analysis;
- visualization of learning resources (Fischer et al., 2020).

Big Data technology in education has been described by Spanish researchers (2019). The education sector is ambitiously embracing personalization of learning, innovation, and improving learning effectiveness. Such approaches may be implemented using Big Data analytics. Information should be reliable and relevant, and the learning process needs to be based on research published in Web of Science, Scopus, ERIC and PsycINFO (Marín-Marín et al., 2019).

Big Data technology ensures the quality of education. To secure this parameter, careful data analysis is applied, including: data collection, cleansing, processing and storage, which also make sure that junk data (such as advertising, roadblocks) are removed. Effective filtering of junk data results in accurate and reliable information. Big Data technology also provides flexibility in education, making it possible to teach a large number of students while keeping them mobile. Big Data technologies use SCORM eLearning standard, NoSQL, MySQL, Clinet, Zookeeper, HMaster, HRegionserver databases (Li & Zhai, 2018).

Use of Massive Open Online Courses (MOOCs) in education for Big Data Analysis was studied by a team of researchers from Spain, Germany, Netherlands (2018). The research suggests that MOOCs are an alternative way to learn by clustering knowledge and large amounts of information. Today they are not widespread, and awareness needs to be raised among instructors (Castaño-Muñoz et al., 2018).

The opportunities offered by Big Data syllabus were studied by US researchers (2017). Ambitious adoption and use of Big Data in education is prompting the academic community to investigate the curricula development criteria. In many curricula, basic Big Data information is broken down into smaller components using interpretations of the instructor's knowledge. In order to effectively implement Big Data into the syllabus, the Big Data instructor needs to make sure that each component of the curriculum is broken down into smaller segments. Such an approach will also provide a clear understanding of the discipline's assessments and resources (Friedman, 2017).

Big Data technology for higher education was studied by Indonesian scholars (2017). Today, learning with Big Data technology requires a cutting-edge data warehouse. Hadoop technology is a powerful Big Data tool. It is a state-of-the-art data warehouse capable of analyzing Big Data at tremendous speed. This technology has fast startup, seamless integration with other services, is easy to use, and is a reliable and secure technology. However, Hadoop cannot provide efficient storage of large numbers of small files. In the course of training, Hadoop technology is used to store files securely (Santoso, 2017).

The use of Big Data in academia was analyzed by Romanian researchers (2015). Analyzing and working with Big Data in education requires the use of cloud technology such as Hadoop (Logica & Magdalena, 2015).

Yet, getting an insight into the application of Big Data technology in the education sector complements and makes more profound the previous research on the problem.

Problem statement

The use of Big Data technology in education makes it possible to secure strong academic achievements among students, as well as upgrade the education system in order to improve available professional knowledge and skills.

The relevance of this study determined its topic – the use of Big Data technology in education. The paper examines the special aspects of using Big Data technology in education. To achieve

the purpose, the following objectives were set: (1) explore students' tools and commitment to Big Data technology; (2) analyze the relationship between commitment to Big Data and students' academic achievements.

The authors were the first to analyze the relationship/correlation between students' commitment to Big Data technology and students' academic achievements. The use of Big Data technology in education makes it highly effective, personalized and affordable.

Methods and sources

Research design and sample

The study took place at Russian Academy of Education (Moscow, Russian Federation), RANEPA (Moscow, Russian Federation), and Al Ain University (Abu Dhabi, UAE). The design and methodology of the study *were developed* jointly *by the authors*. The sample consisted of 356 third-year students, as the academic programs of third-year disciplines allocate 60–85% of the time for independent study (which is the highest figure among all courses). The average age of respondents was 20.57 ± 0.87; boys accounted for 52% and girls – for 48% of respondents.

Research tools

The questionnaire developed by the authors was used to assess Big Data tools. Respondents rated the Big Data technology tools they use most often in their learning process:

(1) Cloud technologies;
(2) Apps;
(3) Massive Open Online Courses (MOOCs);
(4) Digital educational platforms.

Respondents are expected to specify Big Data tool that they use most often in the learning process and rate the instrument ranging from 0 (like it the least) to 10 (like it the most). Respondents were surveyed using Google Forms. The authors sent students a link to the questionnaire, which was active for 1 month (December 2021) and easily accessible from any modern device.

Academic achievements were assessed on a five-point scale, where:

- "1" means that the student demonstrated poor progress in learning;
- "2" means that the student demonstrated mediocre progress in learning;
- "3" means that the student demonstrated sufficient progress in learning;
- "4" means that the student demonstrated good progress in learning;
- "5" means that the student demonstrated excellent progress in learning.

Statistical analysis of data

Microsoft Office Excel was used to analyze statistical data. Quantitative parameters were calculated using the formula (x ± m), where x is the arithmetic mean and m is the standard error. Calculations were performed at a statistical significance (*p*) of 0.05. To identify the correlation between students' commitment to Big Data technology and students' academic achievements, the authors:

(1) conducted correlation analysis to reveal any relationships/correlations;
(2) applied Pearson correlation coefficient (r). When the absolute value of r is less than 0.2, the correlation was very weak. With r less than 0.5, the correlation was weak; with r less than 0.7, the

correlation was medium; with r less than 0.9, the correlation was high; with r exceeding 0.9, the correlation was very high;

(3) conducted multiple regression analysis to establish the dependence of the students' average score on the independent variables potentially affecting the grade.

Research limitations

The study did not include 1–2-year students, as they have different academic programs.

Ethical issues

This research was conducted in accordance with the principles of the Declaration of Helsinki (Ethical Principles for Research Involving Human Subjects). All respondents *were informed* in detail about *the purposes* and methods of the study. All respondents provided written informed consent to participate in the study. Complete anonymity requirements were met. There was no conflict of interest. University bioethics committees authorized the study for the 2021/2022 academic year.

Results

The research findings suggested that when working with Big Data students rely on innovative digital technology, both online and offline (Figure 1).

All survey respondents use apps when learning and analyzing Big Data. 73.03% (260 respondents) use Moodle, 67.13% (239 respondents) use Zoom, 65.17% (232 respondents) use Quizlet, 50.84% (181 respondents) use Skype, and 35.11% (125 respondents) use Slack. The findings suggest that all respondents use apps, which might be attributed to the unlimited chronotope of the learning process, preserving the interaction between students and teachers in formal and informal learning.

MOOCs are used by 75% of respondents. 2.13% (150 respondents) took Becoming a Successful Leader (Leadership Training), 33.99% (121 respondents) – Communication Skills for Dialoguing Across Difference, 28.09% (100 respondents) – Academic and Business Writing, 14.89% (53 respondents) – Receive Your PhD, 13.20% (47 respondents) – Flexible Learning Environment. The findings suggest that 75% of respondents use MOOCs, which might be attributed to the lack of immediate response when a student has a question and does not immediately receive an answer. There is also a lack of personal communication with the instructor.
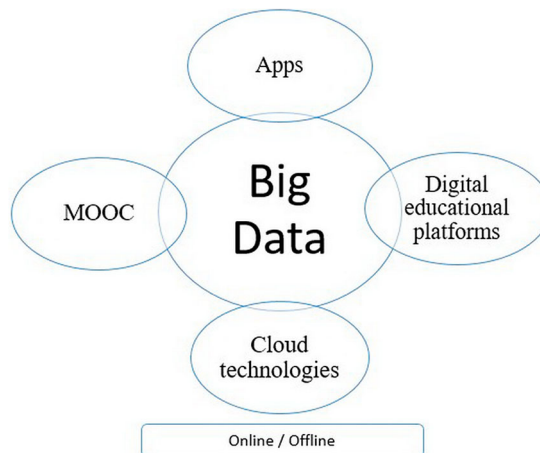

Figure 1. Big Data cutting-edge technology.

Digital learning platforms are used by all respondents. 78.09% (278 respondents) use Microsoft Teams, 51.97% (185 respondents) use Hopin, 36.80% (131 respondents) use Click Meeting, 23.88% (85 respondents) use OpenVidu, and 17.98% (64 respondents) use WebEx (Webex by Cisco). The findings suggest that digital learning platforms are used by all students, which may be due to the mandatory transition of bricks-and-mortar institutions of higher education to innovative formats.

Cloud technology is also used by all students. 96.91% (345 respondents) use Google Drive, 84.83% (302 respondents) use Dropbox, 52.81% (188 respondents) use Amazon Web Services, 51.12% (182 respondents) use OneDrive, 36.24% (129 respondents) use Hadoop. The findings suggest that cloud technology is used by all students, which might be attributed to the ease of use and the fact that, for example, Google Drive has been around since 2012 (Figure 2).

The findings suggest that when working with Big Data, students preferred apps (8.9 ± 1.33) instead of cloud technology (6.9 ± 0.11) (Table 1).

Students believe that important factors affecting the use of Big Data technology in the learning process include:

- quality of information (85.96%, 306 respondents);
- interest in learning (77.81%, 277 respondents);
- instructor support (feedback) (66.85%, 238 respondents);
- time constraints (48.03%, 171 respondents);
- lack of technical skills (31.74%, 113 respondents);
- failure to understand data (10.96%, 39 respondents).

The students' average score was 4.02 ± 0.16.

A multiple regression model was constructed based on the research findings. The dependent variable (Y) was taken as the average student score parameter. The following factors were taken as independent variables, potentially influencing the student grades: use of apps (X1); MOOCs (X2); digital learning platforms (X3); and cloud technology (X4).

When calculating the correlation coefficient, a very weak relationship strength measure (r) was established for factors X3 and X4; a medium relationship was established for factors X1 and X2 (Table 2).
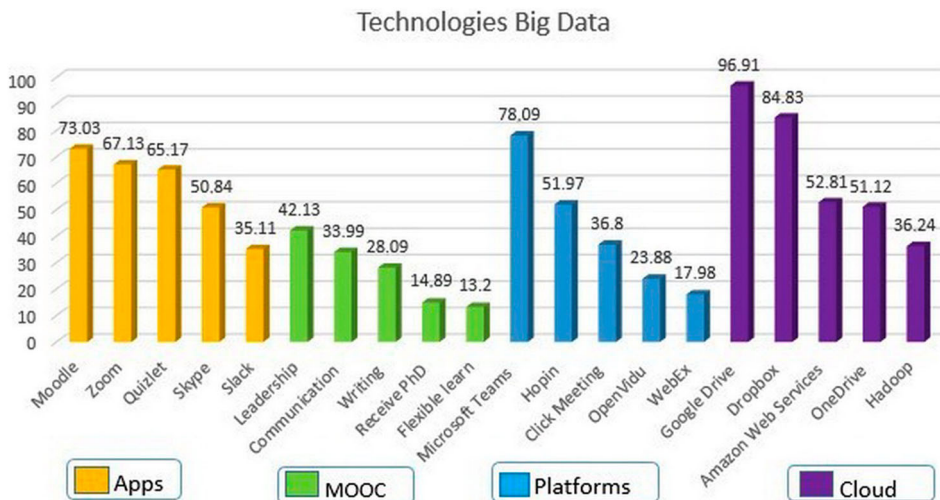

Figure 2. Students' commitment to Big Data technology.

Table 1. The range Big Data tools in education.

No.	Big Data tools in education	Evaluation range 0–10, M ± SD
1	Apps	8.9 ± 1.33
2	Massive open online courses	7.6 ± 1.21
3	Digital learning platforms	7.3 ± 0.67
4	Cloud technology	6.9 ± 0.11

p-value: $p1–2 > 0.05$, $p1–3 < 0.05$, $p1–4 < 0.05$, $p2–3 > 0.05$, $p2–4 > 0.05$, $p3–4 > 0.05$.

Table 2. Correlation of Big Data tools.

	Big Data tools		Correlation coefficients, r
1.	Apps	X1	0.389
2.	MOOCs	X2	0.475
3.	Digital learning platforms	X3	0.146
4.	Cloud technology	X4	0.211

The correlation analysis established an average positive correlation for the following Big Data tools: Apps and MOOCs.

Discussion

In education, big data is used with artificial intelligence to create an artificial neural network. The results of validated experiments show that the use of big data in the educational process has a positive effect on the efficiency and productivity of higher education (Ashaari et al., 2021). In addition, big data is used in the context of the University 4.0 model. Based on an experiment conducted in Malaysian universities, it was concluded that a teacher who uses big data in education needs three abilities – managerial, technological and human (Ashaari et al., 2020a). The use of big data is considered in the context of continuing education in order to improve the quality and efficiency of the educational process at the university. The possibility of using big data together with a new approach to learning data-driven decision-making (DDDM) is considered. Malaysian universities conducted an experiment on the possibility of the university joining Big Data Analytics (BDATC) based on connectivity, compatibility and modularity parameters (Ashaari et al., 2020b). In this context, resource-based theory based on data-driven decision-making (DDDM) has been developed to improve the efficiency and productivity of higher education.

E-learning and digital technology in educational institutions have been studied by Australian researchers (Sandu et al., 2021). Zoom is a state-of-the-art technology that keeps the learning process going and makes it effective. A case study in Australia revealed that:

- 72.1% of respondents believe digital learning during the COVID-19 pandemic is the best solution;
- 87.35% of respondents are satisfied with the video and sound quality;
- 63% of the respondents widely use and have extensive experience with Zoom.

The survey revealed that 67.13% of respondents use Zoom in the learning process.

Big Data and digital analytics in university education were studied by Russian researchers (2021) (Ogurtsova & Fadeev, 2021). The sources of Big Data in universities include:

(1) personal data;
(2) data on student exposure to digital learning systems (digital textbooks, online courses);
(3) performance data (learning outcomes);
(4) administrative data (attendance, absences).

The University of Sydney predicts academic achievements based on 77 variables (integration into the learning process, gender, socio-economic status, cultural background, and others). The study established a very weak quantitative correlation between the average student score and the use of digital learning platforms/cloud technology. The average correlation was estimated for apps and MOOCs.

Modeling online learning using Big Data technology was investigated by researchers in Morocco (Dahdouh et al., 2020). Rapid digital transformations, new information and communication technology require Big Data processing options. Moodle and Hadoop platforms are used in education. 73.03% of respondents also use Moodle; 36.24% of respondents use Hadoop.

Systemic learning using digital platforms has been studied by researchers from China (Xu et al., 2020). Effective learning requires self-regulated learning and Big Data technology. The Chinese researchers believe that the state-of-the-art learning infrastructure includes Moodle, Joomla, Dropbox, Google Doc, Quizlet, TEDEd, Vimeo, Youtube, Turnitin, BigBlueButton, Zoom, Panapto, Trello, Slack. The research findings suggest that Big Data technology was used by all surveyed respondents. 73.03% of respondents use Moodle; 67.13% use Zoom; 65.17% use Quizlet; 35.11% use Slack.

US researchers analyzed Big Data and the opportunities cloud (Muniswamaiah et al., 2019). Cloud technology (servers, storage, databases, networks, software, and analytics) facilitate innovation, resource flexibility, parallel processing of Big Data, and are huge in scope. Hadoop tools and techniques are being widely used, making it possible to process large amounts of data in various formats. 36.24% of respondents use Hadoop.

The issue of Big Data in contemporary education has been studied by researchers from Macedonia (Nikolovska et al., 2018). Big Data analytics include the large volume, high speed and variety of data. The researchers used Moodle and Hadoop in order to arrange and optimize the learning process using Big Data. 73.03% of respondents also use Moodle; 36.24% of respondents use Hadoop.

Big Data and Cloud have been studied by researchers in Egypt (El-Seoud et al., 2017). Big Data technology is used to make informed decisions, streamline processes and develop new models. Cloud technology is used to host Big Data, involving consolidation and pooling of resources. Hadoop is an example of a state-of-the-art modern cloud technology. Hadoop is used by 36.24% of respondents. Big data are being widely integrated with cloud technology in education, making it possible to work with almost unlimited data, meeting the privacy and security requirements.

The use of Big Data in education has been studied by researchers in China (Liang et al., 2016). In today's education, working with Big Data involves the use of data mining and learning analytics, making it possible to predict the integration of MOOCs in a learning environment. 75% of respondents also use MOOCs.

The Big Data technology contributes to streamlining of the learning processes by:

- improving learning outcomes;
- accelerating processing reliable and meaningful data;
- personalizing the learning process;
- being an affordable technology.

Conclusions

The use of Big Data technology in education not only enables processing of huge amounts of data, but also provides a high rate of accumulation and processing of arrays of diverse data; in this case, the information used is reliable and meaningful.

The population was made up of 356 third-year students; respondents' mean age was 20.57 ± 0.87; boys made up 52%, girls – 48%. A questionnaire was used to study Big Data technology, where respondents rated cloud technology, apps, massive open online courses and digital learning platforms.

The study suggested that the education sector is ambitiously applying Big Data technology, both online and offline. All surveyed respondents use apps in Big Data learning and analysis: 73.03% use Moodle, 67.13% use Zoom, 65.17% use Quizlet, 50.84% use Skype, and 35.11% use Slack. MOOCs are used by 75% of respondents. 2.13% took Becoming a Successful Leader (Leadership Training), 28.09% – Communication Skills for Dialoguing Across Difference, 13.20% – Academic and Business Writing, 14.89% – Receive Your PhD, 13.20% – Flexible Learning Environment. Digital learning platforms are used by all respondents (78.09% use Microsoft Teams, 51.97% use Hopin, 36.80% use Click Meeting, 23.88% use OpenVidu, and 17.98% use WebEx). All students also use cloud technology (96.91% use Google Drive, 84.83% use Dropbox, 52.81% use Amazon Web Services, 51.12% use OneDrive, and 36.24% use Hadoop). The research findings also suggested that when dealing with Big Data technologies, students preferred apps (8.9 ± 1.33) instead of the cloud (6.9 ± 0.11). Students believe that important factors affecting the use of Big Data technology in the learning process include: quality of the information (85.96%); interest (77.81%); instructor support (66.85%); time constraints (48.03%); lack of technical skills (31.74%); failure to understand data (10.96%). The correlation analysis established a medium positive correlation for Big Data Apps technology and MOOCs, and a very weak correlation for digital learning platforms and cloud technology. The research findings integrate Big Data technology into the learning process, thus improving learning outcomes and providing greater speed in processing reliable and meaningful data. Big Data Analytics online course is expected to be introduced in educational institutions in order to build and improve the knowledge and skills of working with Big Data among potential students.

Acknowledgments

Disclosure statement

Funding

Data availability

Data will be available on request.

References

Ashaari, M. A., Amran, A., Ahmad, N. H., Bakri, H., & Nazri, S. (2020b). Big data analytics technology capability and data-driven decision making in Malaysian higher education institutions: A conceptual framework. *IOP Conference Series: Materials Science and Engineering*, *874*(1), 012021. https://doi.org/10.1088/1757-899X/874/1/012021

Ashaari, M. A., Amran, A., & Singh, K. S. D. (2020a). Big data analytics capability for improved performance of higher education institutions in the era of IR 4.0: A proposed conceptual framework. *International Journal of Innovation, Creativity and Change*, *13*(8), 73–93.

Ashaari, M. A., Singh, K. S. D., Abbasi, G. A., Amran, A., & Liebana-Cabanillas, F. J. (2021). Big data analytics capability for improved performance of higher education institutions in the Era of IR 4.0: A multi-analytical SEM & ANN perspective. *Technological Forecasting and Social Change*, *173*, 121119. https://doi.org/10.1016/j.techfore.2021.121119

Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2020). Big data in education: A state of the art, limitations, and future research directions. *International Journal of Educational Technology in Higher Education*, *17*(1), 44. https://doi.org/10.1186/s41239-020-00223-0

Bawden, D., & Robinson, L. (2020). Information overload: An introduction. In D. Bawden, & L. Robinson (Eds.), *Oxford research encyclopedia of politics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190228637.013.1360

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, *25*(6), 551–575. https://doi.org/10.1080/0969594X.2018.1441807

Castaño-Muñoz, J., Kalz, M., Kreijns, K., & Punie, Y. (2018). Who is taking MOOCs for teachers' professional development on the use of ICT? A cross-sectional study from Spain. *Technology, Pedagogy and Education*, *27*(5), 607–624. https://doi.org/10.1080/1475939x.2018.1528997

Chen, N.-S., Yin, C., Isaias, P., & Psotka, J. (2020). Educational Big data: Extracting meaning from data for smart education. *Interactive Learning Environments*, *28*(2), 142–147. https://doi.org/10.1080/10494820.2019.1635395

Dahdouh, K., Dakkak, A., Oughdir, L., & Ibriz, A. (2020). Improving online education using Big data technologies. In K. Dahdouh, A. Dakkak, L. Oughdir, & A. Ibriz (Eds.), *The role of technology in education* (Chap. 2). Books on Demand. https://doi.org/10.5772/intechopen.88463

Dresner Advisory Services. (2017). Big data adoption: State of the market. *ZoomData*. https://www.zoomdata.com/master-class/state-market/big-data-adoption

El-Seoud, S. A., El-Sofany, H. F., Abdelfattah, M. A. F., & Mohamed, R. (2017). Big data and cloud computing: Trends and challenges. *International Journal of Interactive Mobile Technologies*, *11*(2), 34–52. https://doi.org/10.3991/ijim.v11i2.6561

Ershova, I.V., Elkova, E. E., Petrakov, A. U., Laptev, V. A., & Fiofanova, O. A. (2021). Architecture of university educational processes in the COVID-19 pandemic: From small data to big data and data science. In I. V. Ershova, E. E. Elkova, A. U. Petrakov, Laptev, V. A & O. A. Fiofanova (Eds.), *Advances in natural human-made and coupled human-natural systems research*. Lecture notes in networks and systems. https://doi.org/10.1007/978-3-030-75483-9_126

Fiofanova, O. A. (2021a). Data analysis competencies in professional standards: From data-experts to evidence-based education. In O. A. Fiofanova (Ed.), *Advances in natural, human-made, and coupled human-natural systems research*. Lecture notes in networks and systems. https://doi.org/10.1007/978-3-030-75483-9_127

Fiofanova, O. A. (2021b). Data architecture on digital educational platforms and data-competence of teachers. *Revista on Line de Política e Gestão Educacional*, *1*, 1700–1716. https://doi.org/10.22633/rpge.v25iesp.3.15591

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., Slater, S., Baker, R., & Warschauer, M. (2020). Mining Big data in education: Affordances and challenges. *Review of Research in Education*, *44*(1), 130–160. https://doi.org/10.3102/0091732X20903304

Friedman, A. (2017). Measuring the promise of Big data syllabi. *Technology, Pedagogy and Education*, *27*(2), 135–148. https://doi.org/10.1080/1475939x.2017.1408490

Holland, A. A. (2019). Effective principles of informal online learning design: A theory-building metasynthesis of qualitative research. *Computers & Education*, *128*, 214–226. https://doi.org/10.1016/j.compedu.2018.09.026

Kalaian, S. A., Kasim, R. M., & Kasim, N. R. (2019). Descriptive and predictive analytical methods for Big data. In Kalaian (Ed.), *Web services: Concepts, methodologies, tools, and applications* (pp. 314–331). IGI Global. https://doi.org/10.4018/978-1-5225-7501-6.ch018.

Li, Y., & Zhai, X. (2018). Review and prospect of modern education using Big data. *Procedia Computer Science*, *129*, 341–347. https://doi.org/10.1016/j.procs.2018.03.085

Liang, J., Yang, J., Wu, Y., Li, C., & Zheng, L. (2016). Big data application in education: Dropout prediction in edx MOOCs. In Liang (Ed.), *2016 IEEE second international conference on multimedia Big data (BigMM)* (pp. 440–443). IEEE. https://doi.org/10.1109/bigmm.2016.70

Logica, B., & Magdalena, R. (2015). Using Big data in the academic environment. *Procedia Economics and Finance*, *33*, 277–286. https://doi.org/10.1016/s2212-5671(15)01712-8

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C.-C. (2020). Challenges and future directions of Big data and artificial intelligence in education. *Frontiers in Psychology*, *11*, 2748. https://doi.org/10.3389/fpsyg.2020.580820

Mamedova, G. A., Zeynalova, L. A., & Melikova, R. T. (2017). Big data technologies in e-learning. *Open Education*, *6*(6), 41–48. https://doi.org/10.21686/1818-4243-2017-6-41-48

Marín-Marín, J.-A., López-Belmonte, J., Fernández-Campoy, J.-M., & Romero-Rodríguez, J.-M. (2019). Big data in education. A bibliometric review. *Social Sciences*, *8*(8), 223. https://doi.org/10.3390/socsci8080223

Miah, S. J., Miah, M., & Shen, J. (2020). Editorial note: Learning management systems and Big data technologies for higher education. *Education and Information Technologies*, *25*(2), 725–730. https://doi.org/10.1007/s10639-020-10129-z

Moharm, K., & Eltahan, M. (2020). The role of big data in improving E-learning transition. *IOP Conference Series: Materials Science and Engineering*, *885*(1), 012003. https://doi.org/10.1088/1757-899x/885/1/012003

Muniswamaiah, M., Agerwala, T., & Tappert, C. (2019). Big data in cloud computing review and opportunities. *International Journal of Computer Science and Information Technology*, *11*(4), 43–57. https://doi.org/10.5121/ijcsit.2019.11404

Nikolovska, A., Velinov, A., Spasov, S., & Zdravev, Z. (2018). Framework for big data analytics of moodle data using Hadoop in the cloud. In Nikolovska (Ed.), *International scientific conference computer Science'2018* (pp. 3–9). Technical University of Sofia. https://www.researchgate.net/publication/327654264_Framework_for_Big_Data_Analytics_of_Moodle_Data_Using_Hadoop_in_the_Cloud

Ogurtsova, E., & Fadeev, R. (2021). Big data and digital analytics in the university education. *Noospheric Studies*, *4*(4), 37–44. https://doi.org/10.46724/NOOS.2021.4.37-44

Oi, M., Yamada, M., Okubo, F., Shimada, A., & Ogata, H. (2017). Reproducibility of findings from educational Big data. In Oi (Ed.), *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 536–537). ACM. https://doi.org/10.1145/3027385.3029445

Qu, J. (2021). Research on mobile learning in a teaching information service system based on a Big data driven environment. *Education and Information Technologies*, *26*(5), 6183–6201. https://doi.org/10.1007/s10639-021-10614-z

Quadir, B., Chen, N.-S., & Isaias, P. (2020). Analyzing the educational goals, problems and techniques used in educational Big data research from 2010 to 2018. *Interactive Learning Environments*, in press. https://doi.org/10.1080/10494820.2020.1712427

Sandu, R., Karim, S., & Kayastha, M. (2021). E-Learning challenges using ZOOM and application of artificial intelligence to improve learning in Australia higher education institutes. In Sandu (Ed.), *International conference e-learning 2021* (pp. 109–114). ICT. https://www.ict-conf.org/wp-content/uploads/2021/07/04_202104C012_Sandu.pdf

Santoso, L. W. (2017). Data warehouse with Big data technology for higher education. *Procedia Computer Science*, *124*, 93–99. https://doi.org/10.1016/j.procs.2017.12.134

Saxena, D., & Lamest, M. (2018). Information overload and coping strategies in the Big data context: Evidence from the hospitality sector. *Journal of Information Science*, *44*(3), 287–297. https://doi.org/10.1177/0165551517693712

Shirinkina, E. V. (2020). Learning platforms in the digital transformation context. *Fundamental Issues of Reliability and Quality (Reliability and Quality of Complex Systems)*, *1*(29), 42–48. https://doi.org/10.21685/2307-4205-2020-1-5

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. https://doi.org/10.1016/j.chb.2018.07.002

Xu, X., Zhu, X., & Chan, F. M. (2020). System design of pintrich`s SRL in a supervised-PLE platform: A pilot test in higher education. *Interactive Learning Environments*, in press. https://doi.org/10.1080/10494820.2020.1802296

Zheng, M., & Bender, D. (2019). Evaluating outcomes of computer-based classroom testing: Student acceptance and impact on learning and exam performance. *Medical Teacher*, *41*(1), 75–82. https://doi.org/10.1080/0142159X.2018.1441984