Issues in Information Systems
Volume 22, Issue 4, pp. 320-333, 2021

DOI: https://doi.org/10.48009/4_iis_2021_346-359

Big data analytics in higher education

Dušan Lesjak, *University of Primorska and ISSBS, dusan.lesjak@guest.arnes.si*
Srečko Natek, *ISSBS, srecko.natek@issbs.si*
Frederick Kohun, *Robert Morris University kohun@rmu.edu*

Abstract

There are numerous studies of the big data analytics (BDA) in different areas, among them several related to the higher education area. The student-oriented higher education relies on traditional information and communication technology and solutions. The big data analytics in higher education follows that path and expands its usage within higher education. The main purpose of this paper is to analyze and discuss the existing studies of big data analytics in higher education and recognize the main areas of its usage and opportunities, while putting a great stress on the advantages and disadvantages of its implementation.

Keywords: Big data analytics, higher education and institutions, students, professors, learning, teaching

Introduction

The big data analytics (BDA) has shown great potential to create a new space for development of the information industry. Literature reviews discover broad interests, implementation and trends of developing BDA in different areas, including higher education (HE).

Despite of the growing changes happening in the HE area, the role of data helping to address contemporary academic and business challenges is often overlooked. As technologies continue to penetrate all aspects of HE, a surplus of useful 'data traces' is generated. These data can be utilized to support Higher Education Institutions (HEIs) to adapt better in response to changes.

We believe that HEIs make a wide use of the BDA, following the lead of the more advanced HEIs already successfully applying it, especially in advising and directing students in their study path, the labor market and its need for the skills and competencies of graduates playing - in addition to characteristics and student abilities - an increasing role. We are convinced that BDA can as well be used even before the candidates make their study decisions, thus before choosing and enrolling in the desired study program.

The research will focus on BDA in HE, where the following research questions (RQ) are key to its successful implementation:
RQ1: What are the key research areas of BDA in HE?
RQ2: What are the key advantages and disadvantages of BDA in HE?

In the following sections, we focus on some existing studies to recognize the potential opportunities of implementing BDA to support HEIs and explore the main future challenges to improve their performance and management. Further on, we mainly focus on advantages and disadvantages of BDA in HE, and conclude with some main findings and recommendations for further studies.

What is big data analytics and its issues?

Big Data (BD) describes data that is fundamentally too big and moves too fast, thus exceeding the processing capacity of conventional database systems (Manyika *et al*, 2011). Therefore big data is not just more data. It is much data, that is so mixed and unstructured, and it is accumulating so rapidly, that traditional techniques and methodologies do not really work. Therefore it covers innovative techniques and technologies to capture, store, distribute, manage and analyze larger-sized data sets with diverse structures.

Generally, BD has come to be identified by a number of key characteristics, which are (Ben 2015):

- Volume—large amount of information is often challenging to store, process, transfer, analyze, present.
- Velocity—relating to increasing rate at which information flows within an organization (e.g., organizations dealing with financial information have ability to deal with this).
- Veracity refers to biases, noise and abnormality in data. It also looks at how data, stored and meaningfully mined, regarding the analyzed problem. Veracity also covers questions of trust and uncertainty.
- Variety—referring to data in diverse formats, both structured and unstructured.
- Verification—refers to data verification and security.
- Value—most importantly, has the data been utilized to generate value of the insights, benefits and business processes, etc. within an organization?

BDA examines large amounts of data to uncover hidden patterns, correlations and other insights. With today's technology, it's possible to analyze data and get answers and significant insight from it almost immediately – an effort that is substantially slower and less efficient with the more traditional business intelligence solutions (https://www.sas.com/en_us/insights/analytics/big-data-analytics.html#industries).

BDA defined

The BDA is defined as a set of techniques and tools to process and interpret large volumes of data that are generated by the increasing digitalization of content, the greater monitoring of human activities and the spread of the internet of things. It can be used to infer relationships, establish dependencies, and perform perditions of outcomes and behaviors. Types of BDA are (OECD, 2016):

- Data mining that comprises a set of data management technologies, pre-processing (data cleaning) techniques and analytical methods with the aim of discovering information patterns from datasets.
- Profiling techniques seek to identify patterns in the attributes of a particular entity and classify them.
- Business intelligence tools seek to monitor key operational metrics and create standard reports on a regular basis in the interest of informing management decisions.
- Machine learning encompasses the design, development and use of algorithms that execute a given task while simultaneously learning how to improve its performance.
- Visual analytics are tools and techniques that allow data to be effectively observed, interpreted and communicated through (often interactive) charts and figures).

BDA benefits

The concept of BD has been around for years; most organizations now understand that if they capture all the data that streams into their businesses, they shall be able to apply analytics and get significant value from it. The new benefits that BDA brings to the table, however, are speed and efficiency. Whereas a few years ago a business would have gathered information, run analytics and unearthed information that could be used for future decisions, today it is able to identify insights for immediate decisions. The ability to work faster – and stay agile – gives organizations a competitive edge they didn't have before

The benefits of using BDA include:

- Quickly analyzing large amounts of data from different sources, in many different formats and types.
- Rapidly making better-informed decisions for effective strategizing, which can benefit and improve the supply chain, operations and other areas of strategic decision-making.
- Cost savings, which can result from new business process efficiencies and optimizations.
- A better understanding of customer needs, behavior and sentiment, which can lead to better marketing insights, as well as provide information for product development.
- Improved, better risk management strategies that draw from large sample sizes of dat. (https://searchbusinessanalytics.techtarget.com/definition/big-data-analytics).

BD challenges

BD brings huge benefits, but also big challenges, such as new privacy and security concerns, accessibility for business users, and choice of the right solutions for business needs. To capitalize on incoming data, organizations will have to address the following:

- Making BD accessible. Collecting and processing data becomes more difficult as the amount of data grows. Organizations must make data easy and convenient for data owners of all skill levels to use.
- Maintaining quality data. With so much data to maintain, organizations are spending more time than ever before scrubbing for duplicates, errors, absences, conflicts, and inconsistencies.
- Keeping data secure. As the amount of data grows, so do privacy and security concerns. Organizations need to strive for compliance and put tight data processes in place before they take advantage of BD.
- Finding the right tools and platforms. New technologies for processing and analyzing BD are developed all the time. Organizations must find the right technology to work within their established ecosystems and address their particular needs. Often, the right solution is also a flexible solution that can accommodate future infrastructure changes (https://www.tableau.com/learn/articles/big-data-analytics).

BD risks

It is important to consider at least some of the risks and related ethical questions raised by use of BD research. As already mentioned in the previous section, reliance on BD collected by powerful institutions or corporations raises significant social concerns. Contrary to the view that sees big and open data as harbingers of democratic social participation in research, the way that scientific research is governed and financed is not challenged by BD. Rather, the increasing commodification and large value attributed to certain kinds of data (e.g., personal data) is associated to an increase in inequality of power and visibility among different nations, segments of the population and scientific communities (O'Neill 2016; Zuboff 2017; D'Ignazio and Klein 2020). The digital gap between those who are not only able to access data, but also use it, is widening, leading from a state of digital divide to a condition of "data divide" (Bezuidenout et al. 2017).

Big data analytics in higher education

HE is a system with many stakeholders, such as study candidates, students and graduates, HEIs and their professors, staff and administration, labor market and unions, governmental agencies and bodies, responsible for HE, etc. HEIs are performing in a competitive, to a certain extend unpredictable and a very complex environment. They are under increasing pressure to respond to national, international and global social and economic changes such as the Western world growing demand for science, technology, engineering and mathematics (STEM) graduates, ensuring that their study programs respond and fulfil labor

market needs by providing employable graduates, and paying special attention to the increased transparency and quality both nationally and globally.

The decisions required for dealing with these challenges, demands and changes are complex, and many are made with no access to needed data sources and related analyses, therefore the needed foundations for good decisions are not available to those assigned to make relevant and timely decisions.

Need for BDA strategies

BDA looks like a promising technology, therefore the HEI administration should adopt the right strategy for BDA implementation in everyday activities in many HEI areas. As a developing area within education, a number of authors have stressed BDA as a tool to tackle some of the key challenges currently facing the HE (Siemens, 2011). It could be claimed that the importance of BD is the capability to recognize useful data and convert it into usable information by recognizing patterns and deviations from patterns. Long and Siemen (2011) indicated that BD is capable of addressing some of the key challenges that HE is currently facing. An OECD (2013) report suggested that it may be the foundation on which HE can both reinvent its business model and bring together the evidence to help make decisions about educational outcomes.

BD decision support

Vast quantity of data is being collected and stored in numerous HEI data sources or data collections. Consequently, there is an extensive availability of various digital libraries, online repositories and their related tools, all of which could be used as means for change in the HE field (Borgman *et al*, 2008, Long & Siemens, 2011; Xu & Recker, 2012, Daniel & Butson, 2013, Oi et al, 2017). BD incorporates the emergent research field of learning analytics (Long & Siemens, 2011), which is already a growing area in education. However, research in learning analytics has initially been limited to examining indicators of individual student and class performance. BD brings new opportunities and challenges for HEIs. Long and Siemen (2011) indicated that BD presents the most dramatic framework in efficiently utilizing the vast array of data and ultimately shaping the future of HE. The application of BD in HE was also confirmed by Wagner and Ice (2012), who stated that technological developments have certainly served as facilitator for the increasing presence of analytics in HE. Many others stressed the growing importance of BDA, especially in the area of learning analytics (Daniel & Butson, 2013, Nguyen 2020). Daniel & Butson (2013) even proposed a conceptual framework to describe BD in HE along four components, utilizing the framework as a way to describe and link different data systems: *Institutional analytic* refers to a variety of operational data that can be analyzed to help with effective decisions about making improvements at the institutional level. *Information technology (IT) analytics* covers usage and performance data that helps with monitoring required for developing or deploying technology, developing data standards, tools, processes, organizational synergies and policies. *Academic/program analytics* provides overall information about what is happening in a specific program and how to address performance challenges. *Learning analytics.* The educational data mining community and learning modelling communities have already explored ways to track student behaviors, recording variables such as number of clicks and time spent on a page, and increasingly more nuanced information such as resilience and retention of concepts.

BD Usage, Advantages and Opportunities

Data is stored in so-called student information systems, student social media usage data, learning management systems, student collaboration systems, student Internet of things devices and servers etc., used by students and professors, student libraries and other usage bases, individual computers and other devices and support systems keeping information on students' study pathways and achievements, and study

programs completion rates. When students interact with various digital devices, relevant data can easily be recorded, stored and made available for further analysis, taking into account that they come in different formats of audio, video, text, numbers and pictures. BD in HE also covers data warehouse systems that store large quantities of longitudinal data on students right down to very specific transactions and activities on learning and teaching. As technologies continue to penetrate all facets of HE, valuable information is being generated by students, computer applications and systems (Hrabowski & Suess, 2010). When students interact with learning technologies, they leave data trails that can reveal their sentiments, social connections, intentions and goals. (Ben, 2015) So, BDA could be efficiently used in various stages of student's passage through the HEI study programs, namely during the admission, enrolment, designing and executing study paths, completion and employment (Baer & Campbell 2011, Zimmerman, 2018). In HE, data mining techniques have spurred on powerful movements, among which it is particularly important to highlight learning analytics (Buckingham &Deakin, 2016; Daniel, 2015; Ferguson et al., 2016). BDA could be useful for scanning student entries on course assessments, discussion board entries, blog entries or social media activities, which could produce hundreds of transactions per student per course. These data could be collected in near real time as it is handled and then analyzed to recommend eventual and needed actions. With the arrival of BD, teachers are now able to access student's academic performance, learning patterns and provide instant feedback (Black & Wiliam, 2018). The timely and constructive feedback motivates and satisfies the students, which gives a positive impact on their performance (Zheng & Bender, 2019). We are convinced that the BDA can also be used even before the candidates make their study preferences, namely before choosing and enrolling in the desired study program. The BDA could help HEIs (and potential students) in matching their students' preferences and capabilities and HEI study programs requirements, and by that positively affect efficiency and effectiveness of a higher education system – completion rates and employability of the graduates.

Academic data can help professors to analyze their teaching pedagogy and influence changes in line with student needs, requirement and expectations. Many online educational sites have been designed, and multiple courses based on individual student preferences have been introduced (Holland, 2019). Improvement in the educational sector depends upon acquisition and technology. The large-scale administrative data can play a tremendous role in managing various educational problems (Sorensen, 2018). Therefore, it is essential for professors to understand the value of BD in education in order to diminish educational issues. Moreover, Lnenicka et al. (2020) stressed the need to extend skill base of educators to include an understanding of analytical techniques.

In the HEI context, BD implies the explanation of an extensive range of operational and administrative data collected processes meant for evaluating institutional performance and development in order to envisage upcoming performance and recognize potential issues associated to study programs, teaching and learning, research, etc., and for better decisions and institutional management which improves performance (Hrabowski, Suess & Fritz, 2011a, 2011b; Picciano, 2012, Oblinger, 2012). Chaurasia et al., (2018) believe that BDA will help in better decision making and ensure the predictive analysis in the academic realm, and that further relations can be tested by integrating other constructs like university size and type. Sedkaoui (2018) claims that it changes learning needs and requires a reorientation toward the development of novel approaches and advancements in HEIs, not only in teaching and learning, but also in their management.

BD transparency

Dringus (2012) argues that bringing transparency to learning analytics as a practice could be used to help preventing any wrongful use of data. As the amount of data available for use is ever-increasing, the benefits will come from good learning management, reliable data warehousing, flexible and transparent data mining and extraction, and accurate and responsible reporting. BDA usage raises issues about ethics of data

collection in regard to data quality, privacy, security and ownership. It also raises the question of an institution's responsibility for taking action based on the information available (Jones, 2012). Security and privacy issues are a further challenge to BDA usage in HE. Additionally, risks and security procedures for data protection and privacy are still lacking in many HEIs. Slade and Prinsloo (2013) pointed out that while most HEIs seem to have policies to regulate and govern intellectual property, safeguard data privacy and regulate access to data, these policies might not be adequate to address contemporary challenges associated with BDA in HE. The question of security and privacy in the educational context is widely discussed among the researchers (Chatti et al., 2017; Aguilar, 2018; Pecori, 2018; Kyritsi et al., 2019, Jones, 2019). It is especially challenging to apply appropriate anonymization analytical techniques for releasing data sets without compromising personal privacy. In light of these findings, it is essential that ethical and social issues are seen as a key element of the technical and scientific conditions related with data management and analysis. The ethical management of data is not gained exclusively by regulating the exchange of research and management of personal data nor with the introduction of supervising of research funding, even though these are important strategies. To guarantee that BD are used in the most scientifically and socially forward-thinking way, it is essential to exceed the concept of ethics as something external and unfamiliar to research. An analysis of data science ethical implications should become a basic element of the background and activity of those who take care of data and the methods used to view and analyze it. Ethical evaluations and choices are hidden in every aspect of data management, including those that may seem purely technical (https://plato.stanford.edu/entries/science-big-data/).

Data integration challenges

Data integration challenges are important, especially where data come in both structured and unstructured formats and need to be unified from different sources stored in systems managed by various departments. There is also a huge cost related with gathering, storing and developing algorithms to mine data, a complex and time consuming process. Furthermore, most institutional data systems are not interoperable, so aggregating administrative data, classroom and online data can pose additional challenges (Daniel & Butson, 2013, Daniel 2015). There exist a few expected challenges related to the implementation of BDA in HE. Some of these include challenges associated with getting users to accept BDA as a conduit for adopting new processes and change management (Ben, 2015). Martínez et al (2020) stressed the importance of proper implementation and they recommended two main phases. The first design phase deals with the engagement of professors and pedagogical experts in defining the data and metrics to be used to support the pedagogical concepts. The second consists of an implementation phase including pilots with students and with crucial engagement of professors in commenting their understanding over students' learning processes, and the feedback the professors could offer to them.

Williamson (2018) is claiming that HE systems are beginning to shift toward agile data infrastructure arrangements which will permit new sources and practices of BD to be integrated into organizational and educational processes as well as study programs. Baig et al. (2020) in their study highlight research limitations and depict the future research directions, which could provide guidelines for researchers about what has been explored or need to be explored in the future, such as giving universities a new perception to design mixed programs that merge traditional teaching and learning with web-based teaching and especially learning. Miah et al. (2020) state that emerging analytics technologies have also forced higher education institution for better exploitation of BDA in their processes and activities.

BDA areas of study

The BDA is a fast growing technology being also present in HE since it has a potential to support, enhance, improve and optimize all aspects of not only the study processes but overall HEI activities and processes.

Table 1. Studies on BDA in HE

Area	Authors
Data collections, data bases and warehouses	Borgman *et al.* 2008; Xu & Recker, Long & Siemens, 2012; Wagner and Ice, 2012; Daniel & Butson, 2013; Buckingham & Deakin, 2016; Ferguson et al., 2016; Oi et al, 2017; Nguyen, 2020
Students	Hrabowski & Suess, 2010; Baer & Campbell 2011, Ben, 2015; Zimmerman, 2018; Black & Wiliam, 2018; Zheng & Bender, 2019; Bonderud, 2020
Professors	Sorensen, 2018; Holland, 2019; Lnenicka et al., 2020
HEIs management and performance	Hrabowski, Suess & Fritz, 2011a, 2011b; Picciano, 2012; Oblinger, 2012; Sedkaoui, 2018; Chaurasia et al., 2018
Ethics, transparency, data privacy and security	Dringus, 2012; Jones, 2012, Slade & Prinsloo, 2013; Chatti et al., 2017; Aguilar, 2018; Pecori, 2018; Kyritsi et al., 2019; Jones, 2019
Implementation	Daniel & Butson, 2013, Ben, 2015, Martinez at al., 2020
Limitations and trends	Williamson, 2018; Gimbel, 2018; Baig et al., 2020; Miah et al., 2020

Numerous authors have researched the BDA in HE. This enables BDA of gathered data to improve all aspects of HEIs. Authors focused on various aspects of the BDA in HE, e.g., data collection and data warehouse, data bases, students, professors, academic and business processes, analytics and performance, to a certain extent also to transparency, ethics, data privacy, security and implementation challenges and limitations, as well as trends, as shown in the Table 1. Studies on BDA in HE.

Table 1 shows the main BDA areas, we identified as the main topics, which authors are dealing with in their research and publications, devoted to the BDA in HE, such as data collections, data warehouses and data bases, students and professors, HEIs performance followed by transparency, ethics, data privacy and security, implementation, limitations and trends. Recognized areas are actually the answer to the RQ1: What are the key research areas of BDA in HE?

Discussion and findings

According to the literature review and especially areas as shown in Table 1, which tells us what the majority of authors dealing with the usage of BDA in HE focused on, we defined Table 2 in which we summarize advantages and disadvantages of BDA in HE across defined areas.

Table 2. Advantages and disadvantages of BDA in HE

Advantages of BDA in HE	Disadvantages of BDA in HE
Data collection, data bases and data warehouses	
- Data collections can be centrally stored and administratively supported. - Centralized stored and supported data are proven to be of better quality. - HEIs data are gathered and stored and available to be used for various purposes.	- Stored data is located in different data bases, different warehouses, stored in different formats. - Centralized data sources are less known and understood by potential user or explorers. - Complexity of centralized data sources can distract the potential usage or application.
Students	
- Profiling students to individualize learning activities. - Tailor and personalize study (learning) paths. - Discover hidden and unusual patterns and opportunities for better learning. - Enhance student learning experience and outcomes. - Forecasting student results (exams) progress. - Improve completion rates. - Increase employability of graduates. - Using Machine learning and AI to support HE learning to develop Intelligent Assistant.	- Less student privacy (personal data). - Potential misinterpretation of student activities data. - AI can influence students' perception of student activities and lead to frustration. - Students are exposed to HEI possibly recognizing their behavioral habits, not related to study. - Students can change/adopt their learning activities for avoiding/twist specific BDA results or to fit into HEI's target "behavior patterns", not suitable or efficient for a particular student.

Professors	
- Profiling professors to individualize teaching. - Discover unusual and hidden patterns and opportunities for better and innovative teaching. - Better overview, understanding and control of professors' teaching activities. - Using Machine learning and AI to support HE teaching to develop Intelligent Assistant. - Higher performance of professors' teaching activities and teaching automation.	- Frustrated traditional professors might not want to exploit the BDA advantages. - Pedagogical challenges of looking for ways to personalize teaching. - Focus more on student than on teaching activities. - AI can influence the professors' perception of teaching activities and lead to their frustration. - More tailored and personalized teaching might increase a teaching burden.
HEIs management and performance	
- Better understanding and control of all HEI activities and processes. - Using Machine learning and AI to support HEI management to develop Intelligent Assistant to support academic and support processes. - Improved performance and management of HEIs. - Increased HEI's competitiveness and reputation.	- Orientation more on HEI performance and less on innovation of academic and supporting processes. - Data driven approach on strategy creation based on past data can limit the innovation of strategies and future development. - Policy/economic challenge, addressing the way to optimize educational results at inter/national levels.
Ethics, transparency, data privacy and security	
- Increased transparency of all activities processes. - Enhanced HEI privacy and security management. - Use of algorithms to recognized eventual students' fraud and other unethical activities. - Use of algorithms to recognized eventual professors' fraud and other unethical activities.	- Personal or sensitive data exposure risk. - More concerns regarding data privacy and security. - More opportunities for misuse or unethical use of available data and analyses. - Student to professor confidentiality relationships can be aggravated.
Implementation	
- New knowledge and experience gained. - Improved use of smart application for learning, teaching and administration processes. - BDA implementation enhance digital awareness and changes of organizational culture.	- BDA infrastructure and algorithms implementations are time consuming and costly. - Lack of analytical knowledge of professors, management and administration. - Additional support of BD experts needed.
Limitation and trends	
- Better usage of existing tools. - Deeper understanding of learning, teaching and administration processes accelerate new development cycle and academic opportunities. - Realistic expectations of future developments, achievements, opportunities and challenges.	- BDA and Machine Learning outcome models are difficult to understand and easily misunderstood. - Added value of BDA are limited by the knowledge that is hidden in HEI data and minds. - Increased pace of development of BDA represent an additional burden to HEIs.

The BDA influences HE in various ways. Literature review has mostly recognized advantages of BDA in HE. But the majority of authors also recognized several disadvantages. Thus the research focused on different advantages and disadvantages of the BDA, organized around the key BDA research areas as shown in Table 2: Advantages and disadvantages of BDA in HE. The research areas derived from Table 1: Studies on BDA. As listed in Table 2, there are quite some advantages as well as disadvantages of the DBA in HE.

Area 1, devoted to data sources, collections, data bases and warehouses salutes centrally designed, stored, supported and administered data, which enhances quality and the availability of gathered/stored data for various purposes, including BDA. Further, it draws attention to the fact that centralized stored data is complex, with different meanings, in different formats, is specifically time dependent and less known by potential users, what all can distract potential usage or BDA application to support uniform and synergy view on organizational data,

Area 2 is devoted to students, who can benefit from the BDA by tailored and personalized study paths, which improves the suitability of student's capabilities and other characteristics on the one hand, and study program requirements and labor market expectations on the other, which altogether enhances their learning experiences and outcomes, improves completion rates and increases employability of graduates. On the other hand, BDA can decrease student privacy and produce misinterpretation of students' learning and living behavior, or even misuse of students' personal or sensitive data and analysis for no-HE purpose.

Area 3 talks about professors. It praises the BDA eventual impact on professors' performance through profiling professors based on teaching activities, better overview and control of their activities, innovative teaching and, finally, finding hidden opportunities for better, diverse and innovative teaching. On the other hand, it stresses additional workload by dealing with various and new pedagogical challenges of personalized teaching and eventual rejection of more traditional professors not so IT experienced.

Area 4 is devoted to HEI management and performance and praises deep understanding and control of all HEI activities and processes and improved performance and management of HEIs. New knowledge increases HEI's competitive advantage and accelerates institutional reputation. On the other hand, it stresses the eventual lack of innovation and policy/economic challenges at broader scopes.

Area 5 praises increased transparency of all academic and business activities and processes, enhanced privacy and security management and recognition of eventual fraud or unethical activities of students and as well as professors or administration. On the other hand, it stresses the risks of exposure of personal or sensitive data, challenges and concerns regarding the data privacy and security of all involved. Furthermore, BDA can automate or promote system misuse or unethical use of data and analyses available and aggravated student-professor, student – administration or professors-administration relationship.

Area 6 deals with the BDA implementation, which is one of the most challenging areas for HEIs. New knowledge and experience gained, improved usage of smart apps for the activities and processes, and especially the enhanced digital awareness fostering changes in the organizational changes, are more than welcomed. Yet, on the other hand that causes extra time, devoted to BDA with substantial costs, lack of knowledge and need for additional experts, especially well-respected and experienced data scientists.

Area 7 deals with the BDA limitation and trends. Once we know them, we understand them deeply resulting in more realistic expectations, yet we have to be aware of the complexity and the human and technological limitations of the available knowledge and respect the accelerated development of the BDA tools, approaches and solutions for sustainable HE growth and development,

Having recognized, described and illustrated advantages and disadvantages of the BDA in HE, we also answer to the RQ2: What are the key advantages and disadvantages of BDA in HE?

Conclusion

The practice of advanced HEIs demonstrates the possible use of BDA for their management, not only in advising and directing students in their study paths, with the growing role, besides characteristics - the students' abilities - also of the labor market and its need for the skills and competencies of graduates, but in a much broader sense.

The research focuses on BDA to seek the answer to the RQ1: What are the key research areas of BDA? The main research areas of the BDA can be identified as data collection and storing, students, professors, HEI management and performance, ethics, transparency, data privacy and security, implementation, limitations and trends.

In the final part, the research focuses on advantages and disadvantages of the BDA in HE, seeking the answer to the RQ2: What are the key advantages and disadvantages of BDA in HE? The research pointed out that the BDA has various impacts in/on HE and recognized main advantages and disadvantages of BDA in HE, organized around the key BDA research areas defined by the answer to the RQ1.

The recognition of main BDA research areas, advantages and disadvantages promote HEIs identification of key future challenges to improve overall data-driven decision-making in HE.

The research was mainly limited to literature reviews and focused comprehensive analyses, and comparison of advantages and disadvantages of BDA in HE from various perspectives. The future research should include an empirical survey focusing on whether the defined areas of the BDA are representative enough and potentially identifying the missing or related areas. The future research should also discover the key differences of advantages and disadvantages among main stakeholders in the study process – student, professors, administration, HEI management and external supporters. One last consideration is the issue of mis/dis information (data) in BD. Given the potential bias and negative impact on accuracy in BDA, how can this possible "bad" data be flagged, validated/confirmed or be eliminated from the BD repository. While this issue is not unique to HE and is a potential problem for BDA viability and acceptance in general, it needs to be addressed.

References

Aguilar, S. J. (2018). Learning analytics: At the nexus of big data, digital innovation, and social justice in education. TechTrends,62(1), 37-45.https://doi.org/10.1007/s11528-017-0226-9.

Baer, L. & Campbell, J. (2011). Game changers. EDUCAUSE. Retrieved March 24, 2014, from http://net.educause.edu/ir/library/pdf/pub72034.pdf

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

Ben, K. D. (/2014). Big Data and analytics in higher education: Opportunities and challenges, British Journal of Educational Technology (2014), Volume 46, No. 5.

Ben, K. D. (2015). Big Data and analytics in higher education: Opportunities and Challenges. British Journal of Educational Technology, Volume 46, No. 5.

Bezuidenhout, Louise, Leonelli, Sabina, Kelly, Ann and Rappert, Brian, 2017, "Beyond the Digital Divide: Towards a Situated Approach to Open Data". Science and Public Policy, 44(4): 464–475. doi: 10.1093/scipol/scw036

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.2018.14418

Bonderud, D. (2020). How Higher Ed Can Use Data Analytics to Boost Admissions. EdTech, https://edtechmagazine.com/higher/article/2020/06/how-higher-ed-can-use-data-analytics-boost-admissions

Borgman, C., Abelson, H., Dirks, L., Johnson, R., Koedinger, K., Linn, M. et al (2008). Fostering learning in the networked world: The cyberlearning opportunity and challenge, a 21st century agenda for the national science foundation. Arlington, VA: National Science

Foundation. Retrieved April 18, 2014, from http://www.nsf.gov/pubs/2008/nsf08204/nsf08204.pdf

Buckingham, S., & Deakin, R. (2016). Learning analytics for 21st century competencies. Journal of Learning Analytics, 3(2), 6–21. https://doi.org/10.18608/jla.2016.32.2.

Chatti, M. A., Muslim, A., & Schroeder, U. (2017). Toward an open learning analytics ecosystem. In B. Kei Daniel (Ed.), Big data and learning analytics in higher education, (pp. 195–219). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-06520-5_12.

Chaurasia, S. S., Kodwani, D., Lachhwani, H., & Ketkar, M. A. (2018). Big data academic and learning analytics. International Journal of Educational Management, 32(6), 1099–1117. https://doi.org/10.1108/ijem-08-2017-0199.

D'Ignazio, Catherine and Klein, Lauren F., 2020, Data Feminism,Cambridge,MA: The MIT Press.

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, Proceedings of the International Conference on Educational Technologies, 29 November–1 December, 2013, Kuala Lumpur, Malaysia, 89–96.

Dringus, L. P. (2012). Learning analytics considered harmful. Journal of Asynchronous Learning Networks, 16, 3, 87–100.

Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Vuorikari, R. (2016). Research evidence on the use of learning analytics – Implications for education policy. In R. Vuorikari, & J. Castaño Muñoz (Eds.). Joint Research Centre Science for Policy Report; EUR 28294 EN. https://doi.org/10.2791/955210.

Gimbel, E. (2018). University Leaders Disagree on How to use Data Analytics, Survey Finds.

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.

Hrabowski, F. A. III, Suess, J. & Fritz, J. (2011a). Analytics in institutional transformation. EDUCAUSEREVIEW, pp. 15–28. Retrieved March 24, 2014, from https://net.educause.edu/ir/library/pdf/ERM1150.pdf

Hrabowski, F. A. III, Suess, J. & Fritz, J. (2011b). Assessment and analytics in institutional transformation. Assessment and analytics in institutional transformation, EDUCAUSE Review, 46, 5 (Sept./Oct. 2011). Retrieved August 8, 2013, from http://www.educause.edu/ero/article/assessment-and-analytics-institutional-transformation

Jones, K. M. K. (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. 66:24. https://doi.org/10.1186/s41239-019-0155-0

Jones, S. (2012). Technology review: the possibilities of learning analytics to improve learner-centered decision-making. Community College Enterprise, 18, 1, 89–92.

Kyritsi, K. H., Zorkadis, V., Stavropoulos, E. C., & Verykios, V. S. (2019). The pursuit of patterns in educational data mining as a threat to student privacy. Journal of Interactive Media in Education, 2019(1), 2. https://doi.org/10.5334/jime.502.

Lnenicka, M., Kopackova, H., Machova R. & Komarova, J. (2020). Big and open linked data analytics: a study on changing roles and skills in the higher educational process, International Journal of Educational Technology in Higher Education, 17:28 https://doi.org/10.1186/s41239-020-00208-z

Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. Educause Review, 46(5), 30–40 Retrieved from https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education.

Manyika, J., Chui, M., Brown, B., et al. (2011) Big Data: The Next Frontier for Innovation, Competition, and Productivity. McKinsey Global Institute.

Martinez, C. P. J., Catasús, G. M. & Fontanillas, T. R. (2020). Impact of using learning analytics in asynchronous online discussions in HE, International Journal of Educational Technology in Higher Education. 17:39 https://doi.org/10.1186/s41239-020-00217-y

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

Nguyen, A., Gardner, L., Sheridan, D. (2020). Data Analytics in Higher Education: An Integrated View. Journal of Information Systems Education, Vol. 31(1) Winter 2020.

O'Neill, Cathy, 2016, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy, New York: Crown.

Oblinger, D. G. (2012). Let's talk analytics. EDUCAUSE Review, July/August, 10–13.

OECD (2013), OECD Education at a Glance 2013 OECD indicators, OECD Publishing, Paris.

OECD (2016), OECD Science, Technology and Innovation Outlook 2016, OECD, Paris.

Oi, M., Yamada, M., Okubo, F., Shimada, A., & Ogata, H. (2017). Reproducibility of findings from educational big data. Paper presented at the Proceedings of the Seventh International Learning Analytics & Knowledge Conference (pp. 536–537). ACM: https://doi.org/10.1145/3027385.3029445

Pecori, R. (2018). A virtual learning architecture enhanced by fog computing and big data streams. Future Internet, 10(1), 4. https://doi.org/10.3390/fi10010004

Picciano, A. G. (2012). The evolution of Big data and learning activities in America higher education. Journal of Asynchronous Learning Networks, Vol. 16. Issue 3.

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.

Sedkaoui, S. (2018). How data analytics is changing entrepreneurial opportunities? Intl. Journal of Innovation Science, 10(2), 274–294. https://doi.org/10.1108/IJIS-09-2017-0092.

Siemens, G. (2011). How data and analytics can improve education, July 2011. Retrieved, October 30, 2014, from http://radar.oreilly.com/2011/07/education-data-analytics-learning.html

Slade, S. & Prinsloo, P. (2013). Learning analytics: ethical issues and dilemmas. American Behavioral Scientist, 57, 10, 1509–1528.

Sorensen, L. C. (2018). "Big Data" in Educational Administration: An Application for Predicting School Dropout Risk. Educational Administration Quarterly, 45 (1), 1–93:https://doi.org/10.1177/0013161x18799439

Wagner, E. & Ice, P. (2012). Data changes everything: delivering on the promise of learning analytics in higher education. EDUCAUSE Review, July/August, 33–42.

Williamson, B. (2018). The hidden architecture of higher education: building a big data infrastructure for the smarter university'. International Journal of Educational Technology in Higher Education. 15:12. DOI 10.1186/s41239-018-0094-1

Xu, B. & Recker, M. (2012). Teaching analytics: a clustering and triangulation study of digital library user data. Educational Technology & Society, 15, 3, 103–115.

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

Zimmermann, E. (2018). Q&A: Kandice Porter Explains How Data Analytics Helped Failure Rates Plummet.

Zuboff, Shoshana, 2017, The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power, New York: Public Affairs.