Big Data in Higher Education for Student Behavior Analytics (Big Data-HE-SBA System Architecture)

Kittipong Chinsook[1], Withamon Khajonmote[2], Sununta Klintawon[3], Chaiyan Sakulthai[3], Wicha Leamsakul[3] & Thada Jantakoon[3]

[1] Department of Computer Science, Faculty of Science and Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand

[2] Department of Computer Technology, Faculty of Science and Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand

[3] Department of Information and Communication Technology for Education, Faculty of Science and Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand

Correspondence: Thada Jantakoon, Department of Information and Communication Technology for Education, Faculty of Science and Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand. E-mail: thada.phd@gmail.com

Abstract

Big data is an important part of innovation that has recently attracted a lot of interest from academics and practitioners alike. Given the importance of the education industry, there is a growing trend to investigate the role of big data in this field. Much research has been undertaken to date in order to better understand the use of big data in many sectors for diverse reasons. Big data in higher education, however, still lacks a complete examination. Thus, the purposes of the research were (1) to design the system architecture of big data in higher education for student behavior analytics and (2) to evaluate the system architecture of big data in higher education for student behavior analytics. The research procedure was divided into two phases. The first phase is designing a system architecture for big data in higher education for student behavior analytics, and the second phase is the architecture evaluation by experts. Purposive sampling was used to select ten experts in big data and student behavior analytics. Data collection tools were the system and the assessment of an appropriate model with a five-level rating scale. The statistics used in the data analysis were means and standard deviation. The results showed that the system architecture of big data in higher education for student behavior analytics consists of four elements: a) Big Data Sources for Behavioral Analytics; b) Big Data Sources for Behavioral Analytics Sub-Domains; c) Big data capture and storage for behavioral analytics; and d) big data behavioral analysis. The experts' opinions on the system architecture were at the most appropriate level.

Keywords: big data, higher education, student behavior analytics, learning analytics, Big Data-HE-SBA

1. Introduction

The world is evolving at a breakneck pace as new technologies emerge (Chae, 2019). Individuals nowadays employ a great number of electronic equipment (Shorfuzzaman, Hossain, Nazir, Muhammad, & Alamri, 2019). These gadgets produce a large amount of data every single second (ur Rehman et al., 2019). Current technology and apps are being created to accommodate to this huge data. These technologies and applications can be used to analyze and store data (Kalaian, Kasim, & Kasim, 2019). For researchers, big data has piqued their attention (Anshari, Alas, & Yunus, 2019). Mikalef, Pappas, Krogstie, and Giannakos attempted to define and characterize big data in various ways in 2018. Big Data is a "cultural, technical, and scholarly phenomenon" (Boyd & Crawford, 2012), researchers and analysts alike have struggled to come up with a "rigorous term" (Mayer-Schönberger & Cukier, 2014). Due to different sources of data and knowledge generated by analyzing human activities or information leaks by people, Big Data allows for 'things one can do at a wide scale that cannot be achieved at a smaller scale, to extract new ideas or produce new types of value' (Mayer-Schönberger & Cukier, 2014). This has proved effective in a variety of contexts and higher education also uses analytical research techniques to gather information on student behaviors, processes of learning, and institutional practices. A well-developed concept is needed in learning analytics like big data (Van Barneveld, Arnold, & Campbell,

2012). More generally, it is defined as the evaluation, collection, analysis, and reporting of student data and their contexts to understand and optimize the learning environment (Long & Siemens, 2011). Although emerging learning analytics practices have the potential to change higher education, they are morally ambiguous and raise ethical questions, particularly in terms of student privacy. Since learning analytics often relies on aggregating huge amounts of sensitive and personal student data from a dynamic network of data flows, the issue of whether students have the right to limit data collection practices and express their privacy desires as a means of handling their data and information emerges.

We begin by reviewing student behavior analytics in general. This is accompanied by a discussion of privacy theory, especially when it applies to data management and how certain controls aid and expand human autonomy. We appreciate the importance and limitations of informed consent to our privacy issues. It has historically been the mechanism by which we have tried to control information about ourselves in the age of Big Data. Following that, we'll go through the different instances in which students unintentionally disclose data and knowledge to their company and third parties without having any control over the disclosures. Finally, we propose a Big Data in Higher Education for Student Behavior Analytics System Architecture (Big Data-HE-SBA System Architecture) that strikes a compromise between student and institutional needs in higher education.

2. Literature Review

In the design the System Architecture of Big Data in Higher Education for Student Behavior Analytics, the related studies and literature were as follows:

2.1 Big Data Concept

Social networking outlets and smartphone networks have the most data, but the percentage of usable information is lower as compared to other types of data sources that are more valuable, such as financial and political organizations, academic institutions, and the corporate climate. Big data within the meaning of e-learning systems are referred to information created by trainers, but in particular by students, as defined by the institutions or profession during the training period, and collected through the teaching management systems, multimedia, and social networks. Briggs (2014) described Big Data by four key characteristics. These are Volume, Velocity, Variety, and Value. According to Banica et al., (2014, pp. 5256) the proper description of the term is as follows: "Big data is a massive collection of shareable data originating from any kind of private or public digital sources, which represents on its own a source for ongoing discovery, analysis, and Business Intelligence and Forecasting." By incorporating these features into Big Learning Data, we will be able to better explain the context and significance of each main character using the four Vs approach:

2.1.1 Volume

The size of the results. The limits of Big Data are challenging to define since this is a very relative characteristic within every domain of usage, even schooling. Even if data from multiple students at a single university is used, we still think the big data model is viable if higher education organizations work together to benefit the students and researchers who make use of it.

2.1.2 Velocity

The increasing flow of data necessitates hardware and networking devices capable of carrying more and more information, as well as technological systems capable of processing it as quickly as possible. Big Learning Data would provide students and Instructors with immediate access to information needed in the teaching process, such as correcting a wrong answer on an assessment test, encouraging Instructors to make revisions to course material during class, and answering students' questions in real-time.

2.1.3 Variety

Big Data is a mash-up of all kinds of formats, both unstructured and hierarchical. As a result, Big Learning Data captures, analyzes, and delivers knowledge from diverse contexts to ensure better learning resources; the emphasis is on managing them, so there can be no inconsistent activities or performances.

2.1.4 Value

If Big Data has science or economic merit. So, while businesses must use social media data in combination with internal data to grow their business, the degree of creativity is more important in the educational setting. Big Learning Data's goal is to achieve a high degree of education and awareness, as well as to build programs in research domains that will result in new technologies in all fields.

2.2 Big Data and Higher Education

New opportunities for higher education policy and learning sciences are opening up as a result of the emergence of integrated libraries in data centers. Many learning analytics supporters believe that tracking, archiving, and analyzing student profiles and behaviors would result in greater instructional decision making, improved learning performance for at-risk pupils, enhanced trust in schools due to data disclosure, and significant pedagogical evolutions, among other items (Long & Siemens, 2011). Universities are actively gathering student data to support several learning analytics programs, which we will explore in this segment.

2.3 Learning Analytics and Privacy as Control of One's Data and Information

As universities continue to implement data analytics projects and infrastructures to collect sensitive, detailed student data, the responsibility to do so responsibly will grow. Even when noble and good goals are in sight, such as optimizing learning (however defined), learning analytics approaches to monitor and engage in the lives of students. As a result, learning analytics, like all Big Data methods, is riddled with privacy issues and ethical quandaries that are only growing in scope (Johnson, Adams Becker, Estrada, & Freeman, 2015). The concern then becomes whether those who develop and fund learning analytics programs can provide students with privacy rights. According to the evidence in the literature, learning analytics demonstrate 'blind holes' (Greller & Drachsler, 2012) in educational strategy and 'poses certain additional boundary constraints' (Pardo & Siemens, 2014) surrounding student data and privacy, which could harm learning analysis ability if left unaddressed (Siemens, 2012).

2.4 eAdvising Analytics

Another field ripe for learning analytics is eAdvising programs. The eAdvising framework at Austin Peay State University incorporates a recommendation engine that recommends courses based on students' academic profiles and compares their course direction to the previous experience of peers like them (Denley, 2012). Other eAdvice services alert students of the possibility to complete the courses without returning to a pre-specified courses chart if they are at risk, or of eligible experts giving them priority guidance if students have been considered to 'at risk' (California State University Long Beach, 2014).

2.5 Edge-case Analytics Using Social and Biometric Data

Leading learning analytics thinkers contend that a student's "every click, every Tweet or Facebook status update, every social interaction, and every page read online can leave a digital footprint" (Long & Siemens, 2011, pp 32) that can "make noticeable" previously invisible social learning habits. This 'smorgasbord' (Diaz & Brown, 2012) approach to data aggregation motivates innovative approaches to learning analytics and promotes data 'fishing expeditions' (Mayer-Schönberger & Cukier, 2013) for new perspectives and patterns.

Learning analytics proponents have yet to show the effectiveness of social analytics at scale, although new projects indicate some possible applications. Some institutions are tracking and mining their students' Facebook use (Ho, 2011; Hoover, 2012), while others are scanning RFID chips in student IDs at lecture halls and classrooms to equate attendance with classroom success (Brazy, 2010; O'Connor, 2010). Universities start to consider students' social life, their partnerships, and their networking systems on campuses by monitoring student behaviors with geological information and mapping interpersonal experiences.

In addition, institutions and academics are investigating the use of biometric data in learning analytics. Biometrics for learning analytics proponents contend that measures of a student's pulse rate, body temperature, ambient luminosity, [location and movement], 'among other things, can be useful for recognizing concentration, tension, and sleep cycles, which can assess conditions that hinder or assist learning (Arriba Pérez, Santos, & Rodriguez, 2016). Initial research shows that when biometrics and their analytics are shared with students, such information will allow people to self-govern their attention (Spann, Schaeffer, & Siemens, 2017).

2.6 Comprehensive Profiles

Understanding how diverse groups of students learn is one of the motivating motivations of those who campaign for studying analytic technology. Institutions must create detailed profiles of learners to do this. Businesses that want to achieve the same aim search outward and buy data profiles from data brokers. Higher education institutions examine themselves and mine the wealth of data gleaned from admissions materials and applications.

The details students provide about themselves on admissions applications and supporting materials is not trivial; in particular, it is often sensitive and revealing. Questions about a student's academic performance, such as transcripts and standardized test scores; career ambitions; demographic and socioeconomic information; and family networks, as well as their academic achievement rating, are all included in admission applications. The

ACT (American College Test) and SAT (Scholastic Aptitude Test or Scholastic Assessment Test) records, for example, provide details on the types of events students engaged in during high school, as well as the types of social activities they expect to partake in while in college. Any applications require informative essays about the prospective student's reading patterns and cultural preferences, as well as his or her disciplinary and criminal background. Others might inquire about the student's religion, sexual orientation, or gender identity (Caldwell, 2012; Hoover, 2011; Steinberg, 2010). In general, this data is used to create detailed personal profiles.

Universities lay the groundwork for conducting empirical experiments and making forecasts by developing data-rich student profiles. Institutional players should correlate data profiles of applicants with parts of the current student body to build predictive ratings of the applicant's prospects for achievement, further informing the student enrollment method (Goff & Shaffer, 2014). The technology of learning analysis often correlates digital and analog behaviors of a student with specific sections of their particular profiles as they enter your preferred institution. Since they will compare the digital trials and information available in their particular professions, the efficacy of other apprenticeship analysis applications would be significantly diminished. While admissions applicant data profiles are abundant, they become much richer as other forms of student data are grafted on as students engage with institutional information systems (Jantakun, Jantakun, & Jantakoon, 2021).

The issue is that higher education agencies are unlikely to adequately educate their prospective applicants on how and for whom the personal information they provide on admissions applications would be used. Students clearly expect these applications to guide admissions decisions, but they don't anticipate downstream uses, and universities don't specifically clarify the knowledge practices that rely on this repository of personal data. In reality, applications for admission, the stage at which we would expect universities to determine informed consent, do not even articulate student privacy rights, particularly when it comes to data control; certain organizations even assert a property right to prospective students' records. This approach is particularly troublesome since students may feel compelled to share all of the intimate aspects of their lives because there is always the risk of being refused entry if they do not.

2.7 Some Motivations of Introducing Big Data in E-Learning

Learning management systems (LMS) focused on interconnected shared computing frameworks are used in universities all over the world. Wikis, chat rooms, and blogs empower teachers to track and monitor students' development, and students to interact more effectively among themselves and with their teachers, allowing them to progress faster and more effectively in an information area. The best service for an instructor who needs to know the extent of understanding of the students regarding the subjects suggested for research is resource sharing and exchanging of ideas. As Banica (2014) points out, a discussion of the educational potential of interactive technologies can begin with the perspectives of the interested classes of students on the one hand and learning practitioners on the other.

3. Methods

3.1 Phase 1 System Architecture Design

Design Big Data-HE-SBA System Architecture. Create an instrument for assessing the appropriateness of the system architecture of Big Data in Higher Education for Student Behavior Analytics.

3.2 Phase 2 Evaluates the Appropriateness of System Architecture of Big Data in Higher Education for Student Behavior Analytics

Population: Population is the experts in the field of Big Data and Student Behavior Analytics. Samples groups: Samples are 10 experts in the field of Big Data and Student Behavior Analytics. Chosen by purposive sampling. They are highly experienced experts in these fields for at least 5 years. Variable: Independent variable is the system architecture of Big Data in Higher Education for Student Behavior Analytics. The dependent variable is the appropriateness of the system architecture of Big Data in Higher Education for Student Behavior Analytics. The research instruments were an evaluation of the system architecture of Big Data in Higher Education for Student Behavior Analytics. The statistics were used to collect and look at the data from the questionnaire. The standard deviation and arithmetic mean are computed.

4. Results

4.1 Big Data-HE-SBA System Architecture

System Architecture of Big Data in Higher Education for Student Behavior Analytics(Big Data-HE-SBA System Architecture)
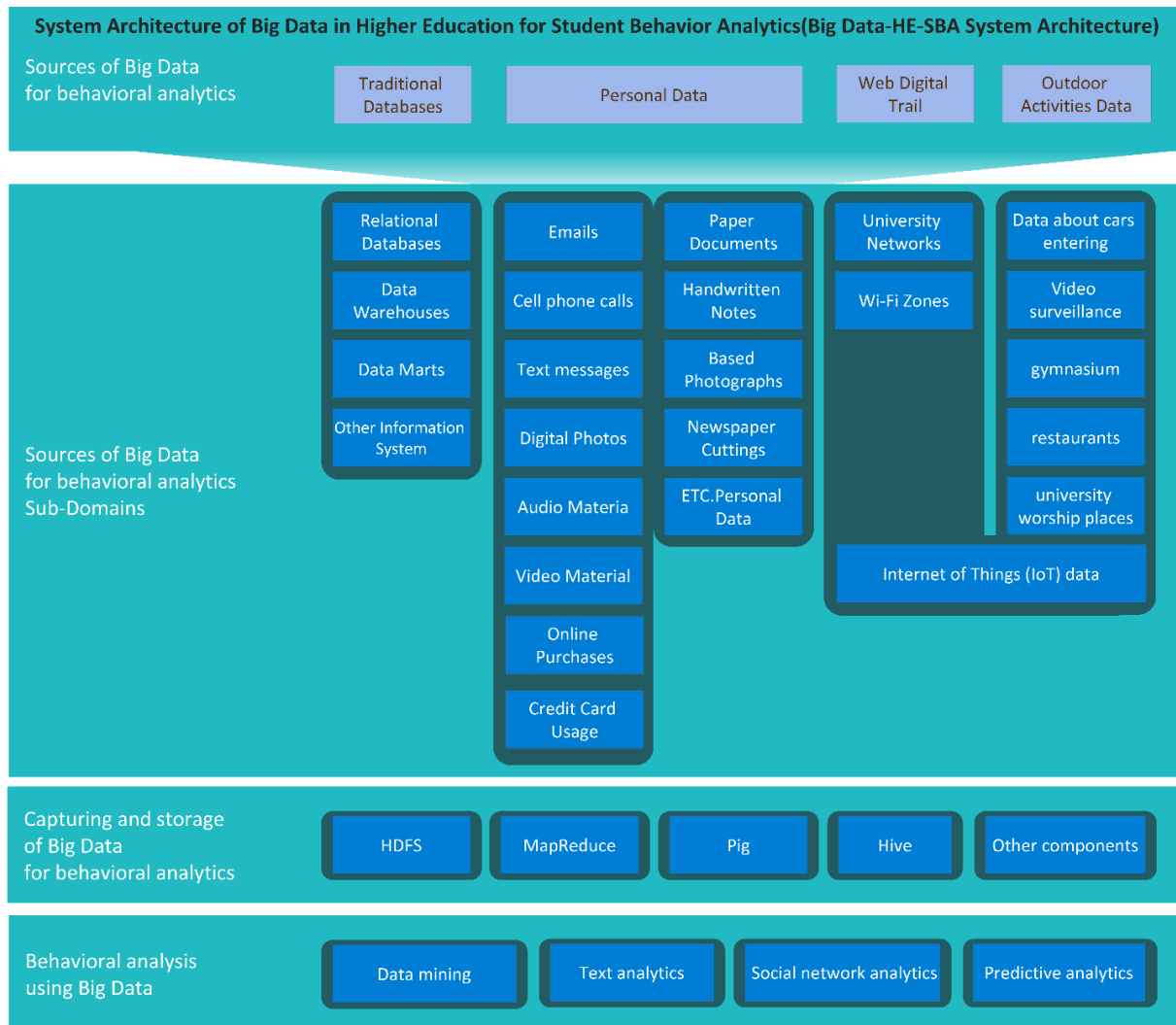
Figure 1. Big Data-HE-SBA System Architecture

4.1.1 Sources of Big Data for Behavioral Analytics and Sub-Domains

Since many variables may affect a student's behavior, behavior analytics is a difficult task. Family, mates, behaviors, and desires are examples of these causes. It's possible that data on these variables isn't readily accessible. Furthermore, any of the available data could be unethical or illegal to obtain. Furthermore, data storage need not be prohibitively expensive or time-consuming. Let's have a look at some of the data that is present in a traditional university that can be used to track and assess a student's behavior.

4.1.1.1 Traditional Databases

Established relational databases, data centers, data marts, and all other software infrastructure producing organized data are traditional data sources. We have details about classes, lessons, exam results, and so on in this list. Databases could be available in the university's hotels, medical facilities, gymnasium, and houses of worship, among other places. Existing files may be supplemented with missing records, such as course schedules, classroom and laboratory assignments, building opening times, teachers' office hours, and so on. We will not only figure out what classes a student is attending, but also where he or she is expected to be at any specific time of day using these databases.

4.1.1.2 Personal Data

This information may be interactive or non-digital. E-mails, phone calls, instant messages, multimedia images, audio and video files, internet orders, and credit card use are all examples of digital records. Paper books, handwritten notes, paper-based photos, newspaper cuttings, and other non-digital evidence are examples. Personal data is normally unstructured or semi-structured, and obtaining it is unlawful or immoral until police have a clear suspicion.

4.1.1.3 Web Digital Trail

A digital footprint is left from all of our daily web-based acts. If students link to the internet via university networks and Wi-Fi zones, their online activities may be tracked. The meta-data of e-mails and a record of pages visited when browsing are two instances. Web and text mining, social network sentiment mining, and data obtained from online portals are some of the more widely employed sources. Social networking networks like Facebook, LinkedIn, and Twitter provide a variety of data that is publicly accessible. Uploading images and videos, leaving notes, sending texts, and pressing the "like" button are all examples of things you may do. Opinion mining, for example, may be used to track certain results. The data is unstructured, and the amount of data access could be insufficient for a person to have useful input. However, with both of the students together, the data could be enormous.

4.1.1.4 Outdoor Activities Data

The university administration is regulated by several outlets, for example.:

- Data about cars entering or leaving a parking area.

-Video surveillance.

-Information from gymnasium, restaurants, and university worship places, etc.

-Internet of Things (IoT).

Surveillance cameras, parking alarms, permission systems for room entry, and a variety of other technologies are often used in the university. Data from these applications is typically tracked and maintained in an isolated way, with local inspection available for identifying security breaches and other rule violations.

Other outlets exist as well, but they might not be compatible with a university-owned scheme. Data on mobile phones pinging cell towers to verify their location and GPS systems tracking a vehicle or a phone are two examples. To find out where they are, mobile phones send pings to cell towers. The IoT is a new trend. In a few years, it is expected to become a common phenomenon.

IoT systems can be cost-effective to install indoors, screens, and electrical switches. The systems would be installed in the building's units and sub-units. This will entail things like a spa, parking, a pharmacy, and restaurants, among other things. We would have apps for automatically turning lights on and off, automated use analysis of rooms and halls, and other things based on these devices. It would be a move toward creating a more energy-efficient and environmentally sustainable climate. Data from IoT devices may be useful for behavioral analytics as well.

4.1.2 Capturing and Storage of Big Data for Behavioral Analytics

Hadoop: It's a system that allows a large number of computers to process a large number of data sets in a distributed manner. Hadoop is a robust framework for dealing with large amounts of data.

4.1.2.1 Hadoop Distributed File System (HDFS)

It's a distributed, modular, and compact file system. It's a method of storing massive files through many servers. As a result, Hadoop may have hundreds or even millions of different files distributed over several devices, all of which are linked by software.

4.1.2.2 MapReduce

It's another crucial component of Hadoop that handles distributed computing. This is a mapping and reducing method. Mapping divides a mission and the associated data into several bits, allowing them to be submitted to multiple servers and processed in parallel. The reducing method blends the outputs from the various computers into a single output.

4.1.2.3 Pig

It is a Hadoop framework for developing MapReduce programs. Pig Latin Programming Language is the language it uses.

4.1.2.4 Hive

Inside Hadoop, there is a data warehouse. It can be used for data summarization, interpretation, and queries. For questions, it employs HiveQL (a SQL-like language).

4.1.2.5 Other Components

There are also other components available. HBase (a NoSQL database), Storm (for streaming data processing), Giraph are some of the most common (used for analyzing social network data), and Spark (for quick in-memory processing).

4.1.3 Behavioral Analysis using Big Data

Several analytical instruments have long been in use. Data analysis, document analytics, network analytics, and predictive analytics are examples of these.

4.1.3.1 Data Mining

Data mining aims to discover unforeseen data trends. Unexpected correlations between variables or individuals clustering together in unexpected ways are examples of these trends.

4.1.2.2 Text Analytics

Text analytics is a form of data mining as well. It is aimed at taking the actual text data content and finding meaning and patterns in words. One of the most important activities of text analytics is sensational research.

4.1.3.3 Social Network Analytics

The analysis of social network dynamics, the identification of important individuals, and the discovery of fascinating patterns of activity are all part of social network analytics. Surprisingly, law enforcement authorities track crime networks as part of this.

4.1.4.4 Predictive Analytics

Predictive analytics employs a variety of methods (such as neural networks, decision trees, and SVM) to attempt to forecast potential occurrences based on historical data.

4.2 The Evaluation of Big Data-HES-SB System Architecture

Table 1. Experts' evaluation of the Big Data-HE-SBA System Architecture

Evaluation Lists	Level of suitability	Level of assessment	
		$\bar{x}$	S.D.
1. Principles and concepts used as the basis for the design of the Big Data-HES-SB System Architecture	Most	4.80	0.4216
2. Big Data-HES-SB System Architecture Processes	Most	4.50	0.5270
3. The Big Data-HES-SB System Architecture	Most	4.65	0.4690
3.1 Sources of Big Data for Behavioural Analytics	Most	4.50	0.5270
3.2 Sources of Big Data for Behavioural Analytics Sub-Domains	Most	4.90	0.3162
3.3 Capturing and storage of Big Data for behavioral analytics	Most	4.60	0.5164
3.4 Behavioural analysis using Big Data	Most	4.60	0.5164
4. The Big Data-HES-SB System Architecture that can be used for practical	Most	4.80	0.4216
Overall	Most	4.65	0.4708

Evaluation of the appropriateness of the Big Data-HE-SBA System Architecture in Table 1 shows that the ten experts agree with the principles and concepts used as the basis for the design of the Big Data-HES-SB System Architecture that had the most level results (mean = 4.80, S.D. = 0.4216). The Big Data-HES-SB System Architecture Processes were the most level (mean = 4.50, S.D. = 0.5270). The Big Data-HES-SB System Architectures that are clear and continuous were the most level (mean = 4.65, S.D. = 0.4690). The Big Data-HES-SB System Architecture that can be used for practical purposes was the most level (mean = 4.80, S.D. = 0.4216), and the overall was the most level (mean = 4.65, S.D. = 0.4708).

5. Conclusions

The composition of the Big Data-HE-SBA System Architecture, derived from the design of four elements of service providers, can be described as follows: a) Behavioral Analytics Big Data Sources; b) Big Data Sources for Behavioral Analytics Sub-Domains; c) Behavioral Analytics Big Data Collection and Storage; and d)

Behavioral Analytics Big Data Analysis. The assessment result of the suitability of the composition of the Big Data-HE-SBA System Architecture from 10 experts reveals that the most level. The results showed that the composition of the system derives from the design, which may be improved by using the Big Data-HE-SBA System Architecture to guide researchers and instructors who want to study, implement, and apply Best Practice to support big data in the higher education management process. Using the latest digital technology, which includes a wide range of tools and techniques that are currently available for the development of big data in higher education for student behavior analytics, such as predictive analytics and data mining, can perform advanced and real-time investigations quickly. Then, use an intelligent teaching system capable of analyzing data from students' interactions during teaching and learning. Implementing higher education platforms, on the other hand, provides various obstacles, especially since it must retain a low level of investment and network management systems to give a wide range of communication and information access. In this research, we presented the design of student behavior analytics for the big data platform. The ideas that underlie the analysis can be used in a wide range of higher education management and instruction, such as improving services to meet the needs of lecturers or students, as well as getting all of the services that are needed for efficient adaptation and deployment in higher education institutions.

References

Anshari, M., Alas, Y., & Yunus, N. (2019). *A survey study of smartphones behavior in Brunei: A proposal of Modelling big data strategies.* In Multigenerational Online Behavior and Media Use: Concepts, Methodologies, Tools, and Applications. IGI global. pp. 201-214. https://doi.org/10.4018/978-1-5225-7909-0.ch012

Arriba Pérez, F., Santos, J. M., & Rodriguez, M. C. (2016). Analytics of biometric data from wearable devices to support teaching and learning activities. *Journal of Information Systems Engineering & Management, 1*(1), 41-54. https://doi.org/10.20897/lectito.201608

Banica, L. (2014). Different hype cycle viewpoints for an E-learning system. *IOSR Journal of Research & Method in Education (IOSRJRME), 4*(5), 88-95. https://doi.org/10.9790/7388-04528895

Banica, L., Paun, V., & Stefan, C. (2014). Big Data leverages Cloud Computing opportunities. *International Journal of Computers & Technology, 13*(12), 5253-5263.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society, 15*(5), 662-679. https://doi.org/10.1080/1369118X.2012.678878

Brazy, D. (2010). *Ariz. College to position sensors to check class attendance*. The Badger Herald. Retrieved from http://badgerherald.com/news/2010/05/04/ariz-college-to-posi/

Briggs, S. (2014). *Big Data in Education: Big Potential or Big Mistake?* Retrieved from http://www.opencolleges.edu.au/informed/features/big-data-big-potentialor-big-mistake/

Caldwell, T. (2012). *More college students may be asked to declare sexual orientation*. The New York Times. Retrieved from http://thechoice.blogs.nytimes.com/2012/03/14/sexual-orientation-university-of-california

California State University Long Beach. (2014). *(EAB) Predictive analytics*. Retrieved from https://web.archive.org/web/20170819031412

Chae, B. K. (2019). A general framework for studying the evolution of the digital innovation ecosystem: The case of big data. *International Journal of Information Management, 45*, 83-94. https://doi.org/10.1016/j.ijinfomgt.2018.10.023

Denley, T. (2012). *Advising by algorithm*. The New York Times. Retrieved from http://www.nytimes.com/interactive/2012/07/18/education/edlife/student-advising-by-algorithm.html

Diaz, V., & Brown, M. (2012). *Learning analytics: A report on the ELI focus session (report ELI paper no. 2)*. Retrieved from https://library.educause.edu/resources/2012/5/learning-analytics-a-report-on-theeli-focus-session

Goff, J. W., & Shaffer, C. M. (2014). Big data's impact on college admission practices and recruitment strategies. In J. E. Lane (Ed.), *Building a smarter university: Big data, innovation, and analytics* (pp. 93-120). Albany: SUNY Press.

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society, 15*(3), 42-57.

Ho, C. (2011). *Companies tracking college athletes' tweets, Facebook posts go after local universities*. The

Washington Post. Retrieved from
http://www.washingtonpost.com/business/capitalbusiness/companies-tracking-college-athletes-tweetsfacebo
ok-posts-go-after-local-universities/2011/10/10/gIQAyHZ9oL_story.html

Hoover, E. (2011). *Elmhurst college will ask applicants about sexual orientation*. The Chronicle of Higher
Education. Retrieved from
http://chronicle.com/blogs/headcount/elmhurst-college-will-ask-applicants-about-sexual-orientation/28553

Jantakun, T., Jantakun, K., & Jantakoon, T. (2021). A Common Framework for Artificial Intelligence in Higher
Education (AAI-HE Model). *International Education Studies, 14*(11), 94-103.
https://doi.org/10.5539/ies.v14n11p94

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2015). *The NMC horizon report: 2015 higher
education edition*. Austin: The New Media Consortium.

Kalaian, S. A., Kasim, R. M., & Kasim, N. R. (2019). *Descriptive and predictive analytical methods for big data*.
In Web Services: Concepts, Methodologies, Tools, and Applications, (pp. 314-331). USA: IGI global.
https://doi.org/10.4018/978-1-5225-7501-6.ch018

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

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and
think*. New York: Houghton Mifflin Harcourt.

Mayer-Schönberger, V., & Cukier, K. (2014). *Learning with big data: The future of education*. New York:
Houghton Mifflin Harcourt.

Mayer-Schönberger, V., & Cukier, K. (2014). *Your high school transcript could haunt you forever: How big data
could create an inescapable "permanent record"*. The Atlantic Retrieved from
https://www.theatlantic.com/education/archive/2014/03/your-high-school-transcript-could-haunt-you-foreve
r/284346/

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: A systematic
literature review and research agenda. *Information Systems and e-Business Management, 16*(3), 547-578.
https://doi.org/10.1007/10257-017-0362-y

O'Connor, M. C. (2010). Northern Arizona University to use existing RFID student cards for attendance tracking.
RFID Journal Retrieved from http://www.rfidjournal.com/articles/view?7628.

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of
Educational Technology, 45*(3), 438-450. https://doi.org/10.1111/bjet.12152

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*(1), 578-588. https://doi.org/10.1016/j.chb.2018.07.002

Siemens, G. (2012). *Learning analytics: Envisioning a research discipline and a domain of practice*. In
Proceedings of the Second International Conference on Learning Analytics and Knowledge, USA. pp. 4-8.
https://doi.org/10.1145/2330601.2330605

Spann, C. A., Schaeffer, J., & Siemens, G. (2017). *Expanding the scope of learning analytics data: Preliminary
findings on attention and self-regulation using wearable technology*. In Proceedings of the Seventh
International Learning Analytics & Knowledge Conference, Canada. pp. 203-207.
https://doi.org/10.1145/3027385.3027427

Steinberg, J. (2010). *University of Pennsylvania tries outreach based on sexual orientation*. The New York
Times. Retrieved from http://thechoice.blogs.nytimes.com/2010/02/26/penn/

Ur Rehman, M. H., Yaqoob, I., Salah, K., Imran, M., Jayaraman, P. P., & Perera, C. (2019). The role of big data
analytics in industrial internet of things. *Future Generation Computer Systems, 92*, 578-588.
https://doi.org/10.1016/j.future.2019.04.020

Van Barneveld, A., Arnold, K. E., Campbell, J. P. (2012). *Analytics in higher education: establishing a common
language* (report no. ELI3026). Retrieved from
https://library.educause.edu/resources/2012/1/analytics-in-higher-education-establishing-a-common-language

Veletsianos, G., Reich, J., & Pasquini, L. A. (2016). The Life Between Big Data Log Events. *AERA Open, 2*(3), 1-45. https://doi.org/10.1177/2332858416657002

Copyrights