# BIG DATA ANALYTICS: AN EDUCATIONAL RESEARCH AGENDA FOR NIGERIA TERTIARY INSTITUTIONS

**[1]UGODULUNWA, C. A. &[2]ESOMONU, N. P. M.**
*[1&2]Department of Educational Foundations, Faculty of Education, Nnamdi Azikiwe University, Awka, Anambra State*
ugodulunwac@gmail.com, npm.esomonu@unizik.edu.ng

**Abstract**
*Big data is an essential aspect of innovation which has recently gained major attention in education. The era of big data demands that traditional data analytics may not be able to handle large volume and variety of data with high velocity. Big data analytics becomes necessary in educational research in order to analyse sets of big data and extract patterns, relationships, and information among educational variables to solve problems in education. It can be deployed efficiently in education because of the emergence of big data technology which involves data mining, data analytics and web dashboards that are useful for facilitating personalized education, improving learning, teaching, and assessment in higher education. Since the major function of higher education is teaching, research and community service, it is imperative that educational researchers adopt methods and tools of big data analytics in collecting, processing, analyzing and reporting of their research findings. The question that arises is how to implement big data analytics in educational research to effectively key into current advances in research. This calls for an educational research agenda in tertiary institutions. This paper begins with a brief introduction where big data and analytics were conceptualized, followed by identification of big data analytics types and models. It also describes the implementation of big data in education and the challenges therein. Finally, the paper concludes with suggestions on educational research agenda on big data analytics for tertiary institutions in Nigeria.*

***Keywords:*** Big data analytics, educational research, big data technology, tertiary institutions

## Introduction

There is quantum of data in all spheres of human life as a result of technological advancement, which has brought the world to the era of big data. Big data is attracting interest in higher education institutions. There is unprecedented growth of data in tertiary institutions arising from technological driven teaching, learning and research. For instance, in Nigeria tertiary institutions, big data are generated from variety of sources. Some of these sources include on-line applications and qualifying examinations by universities, registration data from admitted students, on-line lectures and interactions, teacher-made formative test

scores, and performance test indicators from classroom assessment technique, e-feedback of students' evaluation, and course work results. Other sources include seminars, students' projects, and research outputs from staff and students' internal and external examinations and assessment, as well as enrolment figures, employees' statistics, and financial records (Anikweze, Ugodulunwa & Mainoma, 2019; Esomonu, Esomonu, & Eleje, 2019; Ugodulunwa, et al. 2019).

Globally, higher education institutions operate in a complex and competitive environment and are under pressure to respond to national and global economic, political and social change, such as, the need to assure quality of programmes offered, increase in proportion of students in some disciplines, and embed workplace attributes in their programmes (Daniel, 2014; Tasmin, Muhammad, & Aziati, 2020). In spite of declining government funding and support from private sector, stakeholders expect higher institutions to respond promptly to the needs mentioned above (Thornton, 2013) and also meet regulatory demands for transparency and accountability (Hazelkorn, 2007 as cited in Daniel, 2014). The question that arises is how should higher education institutions respond effectively and efficiently to global changes affecting their environment? Many decisions for handling the rapid changes require recourse to variety of data sources that are generated but not promptly available for use by those responsible for making relevant decisions. Another question to answer is how to develop a high performance platform to analyze big data and how this can be utilized in educational research and mitigate the associated challenges?

The need to discuss and advance an educational research agenda or plan becomes imperative, since big data is still evolving as an important field of research and new findings and tools are emerging. Furthermore, the use of big data analysis techniques for addressing contemporary challenges facing higher education institutions is a current trend that requires attention of the academia (Anikweze, et al., 2019). This is necessary because of the key benefits of big data to organizations, which include improved and timely access to decision-making information, greater transparency, scalability and better change management (Dnuggets, 2018). This paper therefore conceptualized big data and analytics, followed by identification of big data analytics types and models. It also describes the implementation of big data in education and the challenges therein. Finally, the paper concludes by suggesting an educational research agenda on big data analytics in higher education in Nigeria.

**Big Data and Big Data Analytics**

Big data refers to data sets that grow so large and fast that commonly used software tools and storage systems cannot capture, store, manage, and process them within a short time (Elgendu & Elragal, 2014). Mukherjee and Show (2016) describe big data as the gargantuan bulk of data that cannot be dealt with by traditional data-handling techniques. Big data encompasses the use of technologies to capture, process, analyze, and visualize potentially large datasets in a reasonable time frame. Big data is characterized by volume, velocity and variety (Manyika, *et al.,* 2011;

Esomonu, Esomonu, & Eleje, 2019), veracity, value, and verification (Daniel, 2014), as well as vocabulary or context and venue or location (Brands, 2014). According to Tu Zipei, 2012 (as cited in Meng, & Meng, 2014), we call data "big" not only because of its huge capacity but also because of its greater significance in discovering new knowledge and creating new value through exchange, integration and analysis of massive data so that we could get "big knowledge", "big technology", "big profit" and "big development". With the expanding concept of big data, it covers innovative techniques and technologies to capture, store, distribute, manage and analyze larger sized data sets with diverse structures.

Big data analytics (BDA), on the other hand, refers to the process of examining and analyzing large, varied, complex, and dynamic data sets to reveal useful information and improve decision making (Perdhana, 2018). It is defined as the use of advanced data analytic techniques on vast data sets or big data to discover patterns and meaningful use of information (Anirban, 2014; Siemens, 2011). Big data analytics can be described as a process used to extract meaningful insights, such as hidden patterns, unknown correlations from data. It uncovers unapparent relationships in big data sources. The use of big data and data analytics in educational research in the area of educational assessment, curriculum and instruction, institutional administration and management, programme planning, guidance and counseling, resource allocation, among others has great promise for improving access and quality in education.

**Types of Big Data Analytics**

There are different classifications of big data analytics in education sector. Siemens *et al.* (2011) classified big data analytics as learning analytics (LA) and academic analytics (AA). Big data analytics is also classified as educational data mining (EDM) and learning analytics by Linan and Perez (2015). Also, Daniel and Butson (2013) proposed a conceptual framework for describing big data, where data analytics was classified into four, namely institutional analytics, information technology analytics, academic or programme analytics, and learning analytics. The different types of data analytics are briefly conceptualised in this section.

Learning analytics (LA) deals with collection of data about learners and their contexts to for the purpose of understanding and enhancing learning and the environments in which it occurs (Long & Siemen, 2011); while academic analytics (AA) is concerned with improving organizational effectiveness through the use of students, academic, and institutional data (Daniel, 2014). Academic analytics comprises all the activities in higher education institutions that relate to administration, research, resource allocation and management. Educational data mining (EDM) is used to study learners' behaviours and performances. It is also used for analyzing learning processes by investigating learners' interactions with the learning environment through automated interactive learning management system and simulation (Liu, 2014); while institutional analytics is a variety of operational data that can be analyzed for effective decision on how to achieve improvement at

institutional level. According to Daniel and Butson (2013), institutional analytics covers assessment policy analytics, instructional analytics, and structural analytics. Information technology analytics aims at integrating data from different sources such as student information, learning management and alumni systems, and systems managing learning experiences outside the classroom for developing data modeling and analysis (Daniel, 2014).

**Big Data Analytics Models**

Four big data analytics models were as identified by SIMPLilearn (2021). The four models are descriptive, diagnostic, predictive, and prescriptive analytics. Descriptive analytics summarizes past data into a form that people can easily read. It aims at describing and analyzing historical data collected on students, teaching, research, policies, and administrative processes in order to identify patterns from samples and report on current trends such as in student enrolment, graduation rates and progressions into higher degrees (Daniel, 2015). Diagnostic analytics is applied to understand what caused a problem in the first place. Organizations use diagnostic analytics because it provides an in-depth insight into a particular problem. In educational institutions, diagnostic testing which is a powerful procedure for identifying students' learning difficulties for remediation (Esomonu & Eleje, 2020) is a good example of diagnostic analytics.

Predictive analytics looks into the historical and present data to make predictions of the future. It aims at estimating the possibility of future events by looking at trends and identifying associations about related issues and identifying any risks or opportunities in the future (Daniel, 2014). It could be applied on predicting course completion rate, dropout rate or attrition.

Prescriptive analytics prescribes the solution to a particular problem. Perspective analytics works with both descriptive and predictive analytics to assess and determine new ways of operation in order to achieve desirable outcomes.

**Big Data Analytics Models and Tools**

Simplilearn (2021) identified eight stages of big data analytics which he called lifecycle phases of big data analytics. They are business case evaluation, identification of data, data filtering, data extraction, data aggregation, data analysis, visualization of data and final analysis result. In business case evaluation, the big data analytics defines the reason and goal behind the analysis. After which the big data analytics identifies a broad variety of data sources. All identified data from the previous stage is filtered here to remove corrupt data. In data extraction, data that are not compatible with the tool are extracted and then transformed into a compatible form. In data aggregation, data with the same fields across different data sets are integrated. At the data analysis stage, data is evaluated using analytical and statistical tools to discover useful information. The data analytics can produce graphic visualizations of the analysis using different tools. In the last step of the big data

analytics lifecycle, the final results of the analysis are made available to business stakeholders who will take action.

Big data analytics tools include Cassandra, MongoDB, Hadoop, Spark, Talend, and STORM (Simplilearn, 2021). Cassandra is a distributed data base used to handle chunks of data. MongoDB is used on data sets that change frequently. Hadoop is used for storing and processing big data sets, structured or unstructured. Telend is used for data integration and management. Spark is used for real-time processing and analyzing large amounts of data. STORM is an open-source real-time computational system.

**Implementing Big Data Analytics in Education and Challenges**

Emergence of research on big data in education is a recent development and its goal is to examine how to collect and correlate massive volumes of data in order to identify meaningful behavioural patterns and trends rather than storing them (Daniel, 2014). In Nigeria, many universities have learning management systems for managing student information, teaching, learning and assessment, but the data are not correlated to provide feedback for enhancing students' learning and performance.

Currently, answers to assignments and examinations are only measurements on performance of students, although every student generates big data in the form of personal data, social networks, and assessment results, which can be analyzed to obtain optimal learning environment and gain good understanding of students' behaviour. Research on use of big data to guide teaching and learning in Nigerian universities is therefore scarce, if not completely unavailable. Many students are placed on probation and others are withdrawn at the end of each academic year from Nigerian universities because of poor performance. For instance, Opara (2017) reported in the Punch Newspaper that majority of the 700 students rusticated from one federal university in Nigeria in 2015/2016 academic session was due to poor performance. Student attrition in many universities in Nigeria is at an unacceptable level and needs to be addressed. It is therefore assumed that withdrawal of students would be averted if available data being generated yearly are correlated to obtain immediate feedback for students and teachers on how they are performing. The feedback usually comes late after the students have failed, placed on probation or withdrawn.

Big data analytics has different applications in education because it can provide insights into factors that influence success, capture soft elements of learning, which is referred to as deep learning (Thille, et al., 2014), because they are factors within the individual learners, such as, background, behavioural and contextual information. Learning analytics can be used to adjust courses that students take to make them adaptive, flexible, and continually updated to meet current trends in education. The usefulness of big data and learning analytics in the areas of instruction, responsive formative assessment, actively engaged and collaborative learning has been stressed in literature (Cope & Kalantzis, 2016). The benefits

include monitoring students' performance and progress, providing data on instructional interactions, providing longitudinal analysis for predicting students' behaviour pattern, identifying students at risk and alerting teachers to take appropriate action, as well as helping institutions to improve student retention and academic planning (Picciano, 2014; Macfadyen, Dawson, Pardo, & Gasevi, 2014; Bamiah, Brohi & Rad, 2018). At the institutional level, big data analytics can be applied to measure programme performances and foster institutional policy-making based on evidence, while at the inter-institutional level, policy-makers can have a broader view of inter-institutions collaboration opportunities as observed by Perdhana (2018).

Currently big data analytics are being applied in universities in Europe and America for improving instructions, course recommendations, accessibility, research, evaluation, performance prediction, and accountability, among others despite the challenges of implementation (Bamiah, Brohi & Rad, 2018). Since big data and data analytics are permeating processes in higher education, there is need to explore its application in research in the following areas: performance prediction, research and development, behaviour investigation, evaluation and accountability, identification of at risk learners, absence tracking, detection of dropout risk rates, human resource and financial management, among other areas in higher education.

There are some challenges in the application of big data analytics. They include security and ethics (Perdhana, 2018; Bamiah, Brohi & Rad, 2018), cost of infrastructure, data privacy, lack of organization skills and technologies, lack of knowledge of big data and lack of experience in big data use (Alsghaier, Akour, Shehabat & Aldiabat, 2017; Daniel, 2017). Others are getting users to accept big data as a channel for adopting new processes and change management, integration of data from different sources and formats and possible loss of data during cleansing (Daniel, 2014; Daniel, 2017; Tasmin, Muhammad & Aziati, 2020), problem within collaboration between stakeholders (Tasmin, Muhammad & Aziati, 2020), among others.

Efforts are being made to mitigate the challenges because of the numerous benefits of big data analytics in resolving critical issues in higher education institutions. For instance, Daniel (2017) suggests the need to put together a data management strategy and adopt a culture of data-driven decision making in higher education sector, developing national and international standards for data collection, management, security, interoperability, privacy, and access; as well as developing a clear big data institutional strategies that should reflect data governing structures and data utilization policies. These suggestions hold great promise for tertiary institutions in Nigeria whose core objectives are teaching, research and community service to key into big data analytics for solving numerous problems in education, hence the need for a research agenda on big data analytics deployment.

**An Educational Research Agenda on Big Data Analytics**

Big data analytics has come to stay and should be applied to leverage change and enhance decision making in tertiary institutions in Nigeria. In the light of this, there is need for educational researchers to develop a plan of action on big data analytics for solving numerous problems in education. Suggested plan of action is as follows:

1. Conceptualizing, operationalizing, and developing theoretical underpinnings of big data and analytics in tertiary institutions;
2. Investing on big data architecture that is capable of connecting to data sources, data governance, systems management, and protecting quality of service;
3. Setting up a functional data analysis process by creating a community website;
4. Developing a clear institutional big data analytics policy and strategic plan for wide acceptance of analytics in higher education institutions;
5. Collaborating with prominent technology companies to get access to useful teaching and learning applications;
6. Building a new generation of professionals and up-skilling the existing ones with interdisciplinary competences on big data analytics, as well as designing science programmes with big data analytics, and learning science and technology as areas of specialization;
7. Mining data for insights into students' behaviour, learning processes, and institutional practices using learning analytics technology;
8. Using big data analytics in education for research and development, evaluation and accountability, performance prediction, behaviour investigation, tracking of at risk learners, following up graduates, absence tracking, visual analytics on learners' interaction with a forum and different social media;
9. Developing national standards as benchmark for big data collection, management, security, interoperability, privacy, and access, as well as big data analytics application by National Universities Commission (NUC), National Board for Technical Education (NBTE), and National Commission for Colleges of Education (NCCE);
10. Embarking on meta-analysis of studies on big data analytics; among others.

**Conclusion**

This paper examined big data analytics. It also explored possibilities of implementing it in educational research in order to analyze sets of data and extract patterns, relationships, and information among educational variables for solving problems in education. Since research is a major function of higher education, it is imperative that educational researchers adopt big data analytics and tools for collection, processing, analyzing and reporting their research findings. This cannot be achieved without a clear understanding of big data analytics and its applications in tertiary education institutions. Anticipated challenges associated with its implementation can be addressed by establishing an actionable research agenda.

Educational researchers will then focus their research efforts towards the current trend of using big data analytics to enhance students' learning and performance.

**References**


Adegbija, M. V., Fakomogbon M. A., & Daramola, F. O (2012). The new technologies and the conduct of e-examinations: A case study of National Open University of Nigeria, *British Journal of Science, 3*(1), 59-66. Retrieved from http://www.ajournal.co.uk/pdfs/BSvolume3(1)/BSVol.3%20(1)%20Article%203.pdf

Anikweze, C. M., Ugodulunwa, C. A., & Mainoma,H. M. (2019). Survey of the challenges confronting the application of big data in Nigerian higher education. *International Journal of Innovative Research in Education, Technology and Social Strategies, 6*(1), 52-64. Retrieved from http://internationalpolicybrief.org/journals/international-scientific-research-consortium-journals/intl-jrnal-of-innovative-research-in-edu-tech-social-strategies-vol6-no1-march-2019

Anirban, S. (2014). Big data analytics in the education sector: Needs, opportunities and challenges. *International Journal of Research in Computer and Communication Technology, 3*(11), 12-20. Retrieved June 10 2018 from 67a4/28c6764c7ece121dcd@c196f9541c7b2d9f2.pdf

Bamiah, M. A., Brohi, S.N., & Rad, B. B. (2018). Big data technology in education: Advantages, implementations, and challenges. J*ournal of Engineering Science and Technology,* Special Issue on ICCSIT, 229-241.

Brands, K. (2014). TECH practices: Big data and business intelligence for management accountants. *Strategic Finance*. Retrieved fromhttp://sfmagazine.com/wcontent/uploads/sfarchive/2014/06/TECH-PRACTICES-Big--and-Business-Intelligence-for-Management-Accountants.pdf

Chai, W., Labbe, M., & Stedman, C. (2021). *Big data analytics*. https://searchbusinessanalytics.techtarget.com/definition/big-data-analytics

Cope, B., & Kalantzis, M. (2016). *Big data comes to school: Implications for learning, assessment, and research*. Retrieved August 1, 2018 from http://journals.sagepub.com/doi/pdf/10.1177/2332858416641907

Daniel, B. (2014). *Big data and analytics in higher education: Opportunities and challenges. British Journal of Educational Technology, 46*(5), 904-920. doi:10.1111/bjet.12230

Daniel, B. (2017). Contemporary research discourse and issues on big data in higher education. *Educational Technology, 57*(1), 18-22.

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, pp. 89–96.

Dnuggets, K. (2018). *Big data assessment: Key business drivers, expected benefits and common challenges.* https://www.kdnuggets.com/2014/06/big-data-assessment-business-drivers-benefits-challenges.html

Elgendy, N., & Elragal, A. (2014). *Big data analytics: A literature review paper.* Conference paper in lecture notes in Computer Science. Doi:10.1007/978-3-08976-8_16

Esomonu, N.P.M., & Eleje, L.I. (2020). Effect of diagnostic testing on students' achievement in secondary school quantitative economics. *World Journal of Education, 10*(3), 178. DOI:https://doi.org/10.5430/wje.v10n3p178

Esomonu, N. P. M., Esomonu, N. M., & Eleje, L. I. (2019). Assessment big data in Nigeria: Identification, generation and processing in the opinion of the experts. *International Journal of Evaluation and Research in Education (IJERE), 9*(2), 345-351. DOI: 10.11591/ijere.v9i2.20339

Liu, Y. (2014). Big data and predictive business analytics. *Journal of Business Forecasting, 33*(4), 18-21.

Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 30-40.

Macfadyen, L. P., Dawson, S., Pardo, A., & Gasevi, D. (2014). Embracing big data in complex educational systems: The learning analytics imperative and the policy change. *Research and Practice in Assessment, 9*(2),17-28. Retrieved August 4, 2018 from https://eric.ed.gov/?id=EJ1062692

Malik, A. (2021). *Key benefits of big data.* Retrieved from https://bigdata-madesimple.com/benefits-of-bringing-big-data-analytics-into-education/

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxbirgh, C. et al. (2011). *Big data: The next frontier for innovation, competition and productivity.* Retrieved from https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_exec_summary.ashx

Meng, L., & Meng, L. (2014). *Application of big data in higher education.* Second International Conference on Teaching and Computational Science (ICTCS 2014) Published by Atlantis Press.

Mukherjee, S., & Show, R. (2016). Big data: Concepts, applications, challenges and future scope. *International Journal of Advanced Research in Computer and Communication Engineering*, *5*(2), 66-74.

Opara, E. (2017, January 5). FUT Mina expels 700 students. *Punch Newspaper*. Retrieved August 5, 2018 from https://punchng.com/fut-minna-expels-700-students-2/

Perdhana, Y.F. (2018). *Big data analytics in education: Opportunities and challenges.* https://medium.com/@fajarperdhana/big-data-analytics-in-education-opportunities-and-challenges-1688703257ed

Picciano, A. G. (2014). Big data and learning analytics in blended learning environments: Benefits and concerns. Interactive Journal of Artificial Intelligence and Interactive Multimedia, 2(7), 35-43.