International Journal of Database Theory and Application
Vol. 11, No. 4 (2018), pp.31-40
http://dx.doi.org/10.14257/ijdta.2018.11.4.03

Application of Big Data in Music Education and Counselling Research

Nwokenna Edith Nwakaego[1] and Mkpoikanke Sunday Otu[2*]

[1]*Department of Arts Education (Music), University of Nigeria, Nsukka*
[2]*Department of Educational Foundations (Guidance and Counselling), University of Nigeria, Nsukka*
**mkpoikanke.otu@unn.edu.ng*

Abstract

There is no doubt that utilization of big data can transform the operation of music education and counselling research. To this end, the authors of the present study raised a question as to why big data should be applied in music education and counselling research. Qualitative methodology was adopted to find out from literature why big data should be applied in music education and counselling research. The findings generated in the study indicated several reasons why big data is needed in music education and counselling research. Other issues like how to draw sampling in big data, and sources of big data in music education and counselling research, among others were keenly addressed in the study. Recommendation raised was that music education and counselling researchers should be trained and retrained on big data research techniques, methodology, and analysis for proper adoption and implementation of big data in music education and counselling research. Again, it was recommended that stakeholders in music education and counselling disciplines should receive periodical cognitive counselling to enable them adjust to the pains of performing research with big data.

Keywords*: Big data, music education research, counselling research*

1. Introduction

For more than two decades now, big data have existed in numerous forms. Its repositories is often built and managed in corporations with educational, commercial, entertainment, social and other special needs. Historically, big data began in 1990 following the decision by commercial vendors to offer parallel database management systems for data generation [1]. Since then, many architectural backgrounds have been built to ensure the continuous use of big data. For instance, in 1991, Teradata systems emerged as the first to store and analyze big data –1 terabyte. In 2007, Teradata installed the first petabyte based system and up till now Teradata keeps building structured and unstructured data. As observed further by Sabarmathi and Chinnaiyan [1], LexisNexis Group created big data called a "C++"-based distributed file-sharing framework for data storage and query in 2000. The C++ … stores and distributes and unstructured, semi-structured and structured data across multiple servers. Those using C++ system have the chance to build queries in a C++ dialect called ECL which uses an "apply schema on read" technique to deduce the structure of stored data when it is queried. In 2011 a platform called "High-Performance Computing Cluster" (HPCC) was developed to increase the use of big data concept [1].

Still on architectural background, Google developed a process called MapReduce. According to Dean and Ghemawat [2], MapReduce is a programming model and an

Received (July 15, 2018), Review Result (November 2, 2018), Accepted (November 15, 2018)
* Corresponding Author

ISSN: 2005-4270 IJDTA

associated implementation for processing and generating huge data sets or large amount of data. Through the use of MapReduce, queries are divided and spread across parallel nodes and processed in parallel and the results are then gathered and released. The authors further noted that programs written in MapReduce are automatically parallelized and executed on a large cluster of commodity machines. However, limitations were noticed in MapReduce and a programme called "Apache Spark" was developed in 2012 in response to the limitations. Over the years, many architectural programs have been developed to facilitate the use of big data. A more advance one is big data analytics called "5C" (connection, conversion, cyber, cognition, and configuration) [3].

Technologically, there has been series of development in big data process. For instance, techniques for analyzing data, such as A/B testing, machine learning and natural language processing have been developed. Big data technologies, like business intelligence, cloud computing and databases have been established. Also, visualizations, such as charts, graphs and other displays of the data have been established [4]. Other forms of big data are labeled as data cubes or tensors databases such as data mining, internet, multilinear subspace learning, HPC-based infrastructure, distributed databases, search-based applications, massively parallel-processing (MPP) databases, distributed file systems, cloud computing among others [5]. Many other technologies are used for big data. The good thing about the technologies is that there is system high performance, commodity infrastructure, and low cost. Also, there is notable real or near-real time information in big data analytics. It worth noted also that there is provision for virtualization in big data. That is; virtual method is used to gather data. This enables most of the data to remain in place and only to be taken on demand directly from the source systems [6].

2. Brief Concept of Big Data

The term "big data" refers to the data sets which are complex to the extent that usual data-processing application softwares are not enough to handle them. As Snijders, Matzat and Reips [7] observed, big data usually includes data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process data within a tolerable elapsed time. Big data involves search, sharing, transfer, visualization, querying, updating, information privacy data capturing, data storage, data analysis, and data source [8]. Big data can be used to describe advanced data analytics methods that remove value from data to a particular size of data set [9]. According to the author just mentioned, big data value is generated as a result of creating links between pieces of data, about an individual, about individuals in relation to others, about groups of people, or about the structure of information itself. Dedić and Stanier [10] observed that big data has unstructured data nature. Thus, it requires a set of techniques with some forms of incorporation to divulge insights from datasets that are diverse, complex, and of a massive scale.

The characteristics of big data include high volume, velocity, variety, value and veracity to specific technology and analytical methods for its transformation into value [11] [12]. It is characterized by volume in terms of the large quantity of generated and stored data. Thus, if a data is not large enough it cannot be termed as big data. The variety characteristic of big data is determined by its type and nature. Thus, big data is identified in huge text, images, audio, and video. In terms of velocity, big data is determined in terms of the speed the data is produced and processed to meet certain challenges. For data to be termed "Big data" it must often available in real-time. Thus, it must have high velocity ratio. In terms of veracity, it can be noted that big data must maintain high quality and can vary greatly in analysis [13]. Another characteristic of big data is that it is often a cost-free byproduct of digital interaction [14]. On the whole, big data is

characterized by the use of parallel computing tools to hold data [15]. In the present study, big data is described as larger volume, high velocity, and great veracity of data.

3. Sources of Big Data in Music Education and Counselling Research

According to Bhadani and Jothimani [16], the sources of big data are:

3.1. Digitization of content by industries.
3.2. Advancements in technology.
3.3. Use of "smart" instrumentation is another source of big data.
3.4. Social media platforms.
3.5. Internet.
3.6. Media/Entertainment.
3.7. Telecommunications.
3.8. Transportation, logistics, retail and utilities.
3.9. Video Surveillance.

4. Stages of Big Data Process in Music Education and Counselling Research

The following are the stages of big data process in music education and counselling research:

4.1. Data Creation Stage: Creation of data is the first stage in big data process. In music education and counselling research using big data, data can be created from many sources such as data from, music downloads, movie downloads, live watch, call detail records, blogs, Whatsapp, Tweets and Facebook Pages [16].

4.2. Data Collection Stage: This is the stage that the data is collected from various possible data sources. Big data collection methods such as log files, sensors, web crawlers and network monitoring software are recommended for this assignment [17].

4.3. Data Transmission Stage: This is the stage whereby the data collected is transferred to data storage and processing infrastructure. The transmission is done to enable further data processing and analysis. As [16] explained, data transmission can be performed either through Inter-Dynamic Circuit Network (DCN) or Intra-DCN. The former concerns with the transfer of data from the data source to the data center while the latter focuses on the transfer within the data center.

4.5. Data Pre-processing Stage: This is the stage whereby the data collected is pre-processed to improve the data quality needed for analysis. Data pre-processing covers processes such as integration, cleaning and elimination of redundant data.

4.6. Data Storage Stage: This is the stage in which big data storage systems are created to offer reliable storage space and convenient access to the data. Factors such as availability, consistency, and partition tolerance should be considered while creating big data storage. For quick access, the following storage systems are recommended for the storage of big data in music education research: Dynamo, Google File System (GFS), BigTable, Dryad, CouchDB and Cassandra [16].

4.7. Data Analysis Stage: At this stage, the big data that the researcher collected, transformed, pre-processed and stored are analysed using appropriate metrics, architecture, algorithms, tools and visualization.

4.8. Decision Making Stage: At this final stage of the process, decision is made based on the analysis results.

5. Why Should Big Data be applied in Music Education and Counselling Research?

5.1. There is Demand for Big Data in Music Education and Counselling Research: Generally, the demand of information management specialists has increased rapidly over the years because of big data. Accordingly, software AG, Oracle Corporation, IBM, Microsoft, SAP, EMC, HP and Dell spend huge amount of money to specialize in data management and analytics. There is corresponding increase in the world economy due to escalating use of data-intensive technologies. For instance, in a certain report, about 4.6 billion mobile-phone were subscribed worldwide, and between 1 billion and 2 billion people accessed the internet [18]. Another report show that one-third of the worldwide stored information is in the form of alphanumeric text and still image data [1]. Therefore, from the figures, it appears that big data is used in many sectors of the society – government, education, health, music, counselling among others, and many individuals are involved.

5.2. Big Data is Effective in Terms of Productivity, Innovation and Cost in Music Education and Counselling Research: It is observed that the application of big data within music education and counselling permits effectiveness in terms of productivity, innovation and cost. Data analysis often requires multiple parts of music education administrators to work in collaboration and create new and innovative processes to deliver the desired outcome [17,19]. In other words, adoption of big data analysis in music education research offer cost-effective opportunities to improve decision-making in critical development areas [20].

5.3. Big Data Provides Infrastructure for Transparency in Music Education and Counselling Research: Again, big data provides an infrastructure for transparency in music education and counselling research. It has been established that transparency in music education and counselling research requires vast amount of data and advanced prediction tools for a systematic process of data into useful information [21]. Deducing from the authors view points, a prediction in music education research starts with getting hold of data where different type of sensory data is available to acquire such as voltage, pressure, acoustics, vibration, current and controller data. Thus, the generated big data often act as the input into predictive tools and preventive strategies.

5.4. Big Data Helps Improve Healthcare of Music Education and Counselling Stakeholders: In addition, big data helps improve healthcare of music educational and counselling stakeholders –students, parents, teachers, lecturers, and administrators. This is achieved the use of big data analytics which provide personalized medicine, prescriptive analytics, clinical risk intervention, waste and care variability reduction, automated external and internal reporting of patient data, standardized medical terms and patient registries and fragmented point solutions [22]. The healthcare professionals in charge of music education stakeholders use big data in form of electronic health record data, imaging data, patient generated data, sensor data, and other forms of difficult-to-process data [23]. Healthcare data fits into the central scope of big data because broad data in healthcare is now electronic [24]. Therefore, there is need for music education and counselling stakeholders to pay keen attention to health data and information quality online.

5.5. Big Data Broadens the Scope of Music Education and Counselling Research: Furthermore, the application of big data broadens the scope of music education and counselling research. To this end, trained data professionals and managers and a number of universities, have created undergraduates and masters programmes to meet this demand [4,25]. Big data methodology is applied in marketing of music education research. As observed by Wedel and Kannan [26], music education and counselling department of universities may have numerous sub-domains (such as advertising, promotions, product development, training, performance, branding, labeling, etc.) that all may use diverse types of data. Therefore, music education and counselling research should develop ways in which music education and counselling administrators can acquire extensive knowledge on all the techniques used in these sub-domains for accurate application of big data.

5.6. Big Data is used in Music Education and Counselling Media Process: With the use of big data, music education and counselling research can shift away from the conventional approach of using specific media environments such as newspapers, magazines, brochure, television or radio into the use of technologies that reach targeted people at best times in best locations [27]. The aim is to convey a message that is statistically stated to the consumers. Moreover, big data has become a distinctive tool to help employees in music education sector work more efficiently. It has streamlined the collection and distribution of Information Technology (IT) in music education and counselling departments. Again, big data is used to resolve IT and data collection issues in music education and counselling departments.

Usually, big data is gathered in everything that is done online. It is generated with each digital footstep and it's been transmitted in online shopping, social media, streaming movies, downloading music and other transactions [28]. In the context of music education research, big data can change the music education and counselling learners, educators, and administrators' communications. The creation and analysis of big data creates a new paradigm in music education and counselling. The new paradigm is that music education and counselling stakeholders can make enthusiastic choices that enhance learner experiences, streamline teaching practices, keep learners safe, and ensure institutions can account for student attrition, satisfaction, and success. Universities offering music education and counselling courses can gleam demographic, progressive and behavioural data from surveys, tests, and self-assess students' general interaction with online courses.

6. Sampling Big Data in Music Education and Counselling Research

The name big data encloses a term related to size which is a vital feature of big data. But sampling allows the choice of accurate data values from a larger data set to guesstimate the characteristics of the whole population. In music education and counselling research sampling big data can be broken down by various data value groups such as demographic, psychographic, behavioural, and transactional data. Sampling enables the researchers to utilize a customized proportion of population in any study [29]. The authors cited here conducted a study on sampling algorithms for big data. In the study they developed a theoretical formulation for sampling Twitter data. The authors used novel statistical metrics to quantify the statistical representativeness of the Tweet sample, and originate adequate circumstances on the sample size needed for obtaining Tweet samples with high representativeness. In like manner, music education research can utilize big data with accurate sampling. In each case, the population and the sampling will depend on the problem being discussed. For instance, in study on music downloads or direct play on YouTube, the researcher may decide to use the number of clicks per day as population and then apply certain methods to determine the sampling. The rule of the

thumb should be that every big data research in music education and counselling passes through sampling.

7. Examples of Studies that Applied Big Data in Music Education Research

Mcpherson and Davidson [30] collected big data, including broad interviews and surveys with teachers, students and parents. Among others, video data of practice strategies, and assessment of visual and creative forms of musical aptitude and skill were performed for over a period of more than five years. In another instance, the authors further revisit as many of the original participants as possible through online surveys, face-to-face interviews, diaries and visual-auto ethnographic methods to examine present music lives and reflect upon past ones. In each of the instances, big data were collected and the researchers passed through the stages outlined earlier. Mills [31] used national data set to establish a robust picture of relevant aspects of music in schools in England. The author also performed the research on an individual school that has systematically increased the participation in music using strategies that other schools could replicate. The studies were performed using big data processes. In another study using big data procedure, Faulkner, Davidson and McPherson [32] reported numerous decision trees that emerged from mining for knowledge in datasets constructed from the musical journeys, experiences and abilities of 157 young people in Australia from the outset of instrumental tuition in primary school and for the following 12 years. Still through big data procedure, Schäfer, Smukalla and Oelker [33] investigated long-term effects of intense musical experiences on people's way of life.

8. Example of Studies that Applied Big Data in Counselling Research

Based on big data procedure, Gorad, Zalte, Nandi and Nayak [34] developed a web application that helps students studying in high schools to select a course for their career. The system also recommends a career option based on the student's personality trait, interest and their capacity to take up the course.

9. Possible Limitations of Application of Big Data in Music Education and Counselling Research

9.1. The rate of increase in big data is much faster than the existing processing systems [17]. Therefore, more processing systems should be developed to carter for big data in music education research.

9.2. It is possible that the value of big data may decrease over time. Thus, real time or near real time analysis of the data is recommended.

9.3. It is difficult to have good representation of heterogeneous data. For instance, unstructured data like social media, images and videos data which cannot be stored is hard to sample with good representation.

9.4. There are challenges such as inadequate technological infrastructure, economic and human resource scarcity in a developing country like Nigeria. Future study should investigate how the challenges can be controlled.

9.5. Also concerns such as privacy, imperfect methodology, and interoperability issues challenge the application of big data.

9.6. Also use of big data in healthcare of music education stakeholders may raise momentous ethical concerns ranging from risks for individual rights, privacy and autonomy, to transparency and trust.

10. Conclusion and Recommendations

In order to apply big data in music education and counselling, first, data must be processed with advanced statistical tools called "analytics and algorithms" [35]. This makes the data to reveal meaningful information both visible and invisible components. For effectiveness, data creation algorithms need to identify and fix visible and invisible issues such as machine degradation, component wear, among other [21].

The authors recommended that music education and counselling researchers should be trained and retrained on big data research techniques, methodology, and analysis for proper adoption and implementation of big data procedure in music education and counselling research.

Again, the authors that stakeholders in music education and counselling disciplines should receive periodical cognitive counselling to enable them adjust to the reality of performing research with big data.

The authors also recommended that music should be used as therapeutic technique in counselling and psychotherapy. This is primarily because music evokes emotion and memory, brings meaning to life, and motivates individuals to go beyond perceived limits. Also, it appears that music awakens human senses and stimulates human brain structures. Therefore, the use music technique in counselling can facilitate expression and relieve of challenging situations. Again music therapies should be utilize in counselling as a rehabilitative tool for various emotionally distressed outcomes such as depression, mood disturbances, anxiety, nausea, fear, anger, sadness, and pain [36].

The authors also recommended that counsellors should create chance to participate in music-based activities such as drumming, singing, song selection, song writing, movement and dance, music-based imagery, visualization and combined art-music activities. Such participation will give them opportunity to engage in music-based researches to promote music and counselling research collaborations. Also, such participations will give counsellors ample opportunities to learn music skills they can apply in counselling situations.

Acknowledgments

The authors appreciate the previous researchers whose studies are cited in the present study for providing constructive research materials. The authors also appreciate music education and counselling experts who proof read and edited the study. Lastly, the reviewers and editors who reviewed the study are hereby acknowledged for their constructive intellectual suggestions that help improve the standard of this study.

References

[1] A. K. Bhadani and D. "Jothimani, Big Data: Challenges, Opportunities, and Realities, Effective Big Data Management and Opportunities for Implementation", IGI Global, **(2016)**, pp. 1-24.
[3] V. Chang, An overview, examples, and impacts offered by Emerging Services and Analytics in Cloud Computing virtual reality, Neural Computing and Applications, vol. 29, **(2018)**, pp. 1243-1256.
[4] H. Chen, R. H. Chiang and V. C. "Storey, Business intelligence and analytics: from big data to big impact", MIS quarterly (**2012**), pp. 1165-1188.
[5] M. Chen, S. Mao and Y. Liu, "Big data: A survey", Mobile networks and applications, vol. 19, **(2014)**, pp. 171-209.
[6] N. Couldry and J. Turow, "Advertising, big data and the clearance of the public realm: marketers' new approaches to the content subsidy", International Journal of Communication, vol. 8, **(2014)**, pp. 1710-1726.
[7] K. Crawford, "Six provocations for big data", **(2011)**.

[8] A. De Mauro, M. Greco and M. Grimaldi, "A formal definition of Big Data based on its essential features", Library Review, vol. 5 **(2016)**, pp. 122-135.
[9] J. Dean and S. Ghemawat, "MapReduce: simplified data processing on large clusters", Communications of the ACM, vol. 51, **(2008)**, pp. 107-113.
[10] B. N. Ezegbe, M. O. Ede, C. Eseadi, O. O. Nwaubani, I. N. Akaneme, E. N. Aye, K. R. Ede, J. C. Omeje, C. Ezurike and C. N. Onyishi, "Effect of music therapy combined with cognitive restructuring therapy on emotional distress in a sample of Nigerian married couples", Medicine, vol. 97, **(2018)**.
[12] R. Faulkner, J. W. Davidson and G. E. Mcpherson, "The value of data mining in music education research and some findings from its application to a study of instrumental learning during childhood", International Journal of Music Education, vol. 28, **(2010)**, pp. 212-230.
[13] R. J. Goeke, K. A. Crowne and D. R. Laker, "The Effect of Education on Information Systems Success: Lessons from Human Resources", Information Resources Management Journal (IRMJ), vol. 31, **(2018)**, pp. 17-33.
[14] P. B. Goes, "Design science research in top information systems journals, MIS Quarterly": Management Information Systems, vol. 38, **(2014)**, pp. iii-viii.
[15] N. Gorad, I. Zalte, A. Nandi and D. Nayak, "Career Counseling using Data Mining", International Journal of Engineering Science, 10271 (**2017**).
[16] S. Grimes, Big Data: Avoid "Wanna V'Confusion, InformationWeek". Com (**2013**).
[17] V. Huser and J. J. Cimino, "Impending challenges for the use of big data", International Journal of Radiation Oncology• Biology• Physics, vol. 95, **(2016)**, pp. 890-894.
[18] H. Jenkins, S. Ford and J. Green, "Spreadable media: Creating value and meaning in a networked culture", NYU press, (**2018**).
[19] D. Laney, "3D data management: Controlling data volume, velocity and variety", META group research note, vol. 6, (**2001**), pp. 1.
[20] J. Lee, E. Lapira, B. Bagheri and H.-A. Kao, "Recent advances and trends in predictive manufacturing systems in big data environment", Manufacturing Letters, vol. 1, **(2013)**, pp. 38-41.
[21] H. Lu, K. N. Plataniotis and A. N. "Venetsanopoulos, A survey of multilinear subspace learning for tensor data", Pattern Recognition, vol. 44, **(2011)**, pp. 1540-1551.
[22] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh and A. H. Byers, "Big data: The next frontier for innovation", competition, and productivity, **(2011)**.
[23] B. Matturdi, Z. Xianwei, L. Shuai and L. Fuhong, "Big Data security and privacy: A review", China Communications, vol. 11, **(2014)**, pp. 135-145.
[24] G. E. Mcpherson and J. W. Davidson, "Musical practice: Mother and child interactions during the first year of learning an instrument", Music Education Research, vol. 4, **(2002)**, pp. 141-156.
[25] J. Mills, "Using a large data set for research in music education: Music in Years 10 and 11", Bulletin of the Council for Research in Music Education, **(1999)**, pp. 103-107.
[26] T. B. Murdoch and A. S. Detsky, "The inevitable application of big data to health care", Jama, 309 **(2013)**, pp. 1351-1352.
[27] J. O'donoghue and J. Herbert, "Data management within mHealth environments: Patient sensors, mobile devices, and databases", Journal of Data and Information Quality (JDIQ), vol. 4, **(2012)**, pp. 5.
[29] D. S. Palguna, V. Joshi, V. T. Chakaravarthy, R. Kothari and L. V. Subramaniam, "Analysis of Sampling Algorithms for Twitter", IJCAI, **(2015)**, pp. 967-973.
[30] S. Pllana, I. Janciak, P. Brezany and A. Wohrer, "A Survey of the State of the Art in Data Mining and Integration Query Languages, Network-Based Information Systems (NBiS)", 2011 14th International Conference on, IEEE, (**2011**), pp. 341-348.
[31] R. Potts, "Design Education at the Boundary", The Design Journal, 20 (2017), pp. S4263-S4280.
[32] D. J. POWER, "Using 'Big Data'for analytics and decision support", Journal of Decision Systems, vol. 23 **(2014)**, pp. 222-228.
[33] G. Sabarmathi and R. Chinnaiyan, "Investigations on big data features research challenges and applications", Intelligent Computing and Control Systems (ICICCS), 2017 International Conference on, IEEE, (**2017**), pp. 782-786.
[34] S. Sagiroglu and D. Sinanc, "Big data: A review", Collaboration Technologies and Systems (CTS), 2013 International Conference on, IEEE, **(2013)**, pp. 42-47.
[35] T. Schäfer, M. Smukalla and S.-A. Oelker, "How music changes our lives: A qualitative study of the long-term effects of intense musical experiences", Psychology of music, vol. 42, **(2014)**, pp. 525-544.
[36] C. Snijders, U. Matzat and U.-D. Reips, "Big Data": big gaps of knowledge in the field of internet science, International Journal of Internet Science, vol. 7, (**2012**), pp. 1-5.
[37] Y. Wang, R. Goldstone, W. Yu and T. Wang, "Characterization and optimization of memory-resident mapreduce on HPC systems", Parallel and Distributed Processing Symposium, 2014 IEEE 28th International, IEEE, (**2014**), pp. 799-808.
[38] M. Wedel and P. Kannan, "Marketing analytics for data-rich environments", Journal of Marketing, 80 **(2016)**, pp. 97-121.
[39] J. Lee, H. Qiu, G. YU and J. Lin, "Bearing data set. IMS center. university of cincinnati, NASA Ames Prognostics Data Repository".[http://ti. arc. nasa. gov/project/prognostic-data-repository] (**2007**).

[40] G. Pulse, "White Paper: Big Data for Development: Opportunities & Challenges", URL: http://unglobalpulse. org/BigDataforDevWhitePaper (**2012**).
[41] E. Acar, R. J. Harrison, F. Olken, O. Alter, M. Helal, L. Omberg, B. Bader, A. Kennedy, H. Park and Z. BAI, ''Future directions in tensor-based computation and modeling'', NSF Workshop Rep., Arlington, VA, (**2009)**.
[42] N. Dedić and C. Stanier, "Towards differentiating business intelligence, big data, data analytics and knowledge discovery", International Conference on Enterprise Resource Planning Systems, Springer, (**2016)**, pp. 114-122.
[43] J. Lee, B. Bagheri and H.-A. Kao, "Recent advances and trends of cyber-physical systems and big data analytics in industrial informatics", International proceeding of int conference on industrial informatics (INDIN), (**2014**), pp. 1-6.
[44] G. Pangal, M. B. Schmitz, V. Ravindran and E. D. Mcclanahan, "Apparatus and method for data virtualization in a storage processing device", Google Patents, (**2008**).

