Big Data Analytics for Preschool

Mohamed Bellaj[1] [1], Abdelaziz El Hibaoui[1] [2] and Ahmed Bendahmane[2] [3]

[1]*Computer Science and Systems Engineering Laboratory, ABDELMALEK ESSAADI University, Faculty of Science, Tetuan, Morocco*

[2]*Applied Sciences and Education Laboratory, ABDELMALEK ESSAADI University, Higher Normal School, Tetuan, Morocco*

Keywords: Big Data, Preschool, Education; Analytics

Abstract: This paper offers a detailed examination of Big Data Analytics (BDT) benefits, processes, and challenges in the preschool education sector. By simplifying institutions, organizations, instructor techniques, statistical instruction, and evaluation processes BDT plays an essential role in maximizing education intelligence. Furthermore, BDT was used to examine, classify, and forecast learners' needs, as well as vulnerability shortcomings and results, to advance their awareness outcomes and ensure that instructional programs are of high quality. The phases of jobs Big Data and how to process it were also described in this paper. While BDT makes a significant contribution to education, it faces several challenges in terms of security, privacy, ethics, and a shortage of qualified personnel, as well as data processing and storage. As a result, we will discuss the causes of some of the problems associated with applying big data analytics in the preschool market, as well as some recommendations for overcoming those challenges.

1 INTRODUCTION

In the educational sphere, professors' entire pedagogical decisions to assess a student's comprehension of the content or plan the layout of a course can have the most significant impact on student learning and graduation. High-grade lessons can shorten the time it takes a student to learn a particular subject, allow students to obtain more knowledge in the same amount of time, and assist them in making effective choices about what they can specifically learn and drill.

This learning productivity not only improves the student's capacity but also saves time. However, it also benefits teachers by reducing some of their demands. Big Data Analytics is the most practical method for enabling the flexible decision-making that educators need to improve the consistency of the educational environment. It is a cost-effective way to give educators and apprentices an advantage in determining when and how improvements should be made in the learning process (A. Franco, Pablo Daniel, Antonio Matas, and Juan José Leiva. 2020).

A series of decisions, most lauded in the area of product innovation, has been made to modernize the education sector industry. Big Data Analytics opens up new possibilities for advancing the educational process. Assist teachers and students in making more informed choices early in the learning process (Reidenberg, Joel R., and Florian Schaub. 2018). Data science is being used to accelerate practice creativity, and progress is being made quickly. Nearly every day, new technologies and intelligent applications are developed to assist students and educators in making better use of their time.

Technology has always been and will continue to be an important part of our work. Even, the most important thing is how we, as educators, use the influence of digital technologies to help our decisions. What is commonly called “soft skills”; provides teachers with the ability to develop cognitive and emotional intelligence.

[1] https://orcid.org/0000-0002-7057-6107
[2] https://orcid.org/0000-0003-2167-0831
[3] https://orcid.org/0000-0003-3843-4800

Bellaj, M., El Hibaoui, A. and Bendahmane, A.
Big Data Analytics for Preschool.
DOI: 10.5220/0010736500003101
In *Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning (BML 2021)*, pages 457-465
ISBN: 978-989-758-559-3

Educators and students must understand how data analytics can improve the learning experience to create the correct interrogations and to mark the preeminent use of big data as a device to advance our decision-making.

Most states do not systematically collect information on how early childhood education programs collect and use data. Given this lack of information, the results from the current study help provide the early childhood community with detail on data collection and use in early childhood education classrooms (Zweig Clare W. Irwin Janna Fuccillo Kook Josh Cox .2015).

2 BIG DATA AND EDUCATION

Big Data, in its proper sense, refers to the massive amount of data that floats over several stations and, more broadly, digitally in any fleeting moment. It's data that's too large, complex, and volatile for traditional tools to release and handle. The format emerged in the free public domain, where experts were attempting to find faster and more usable ways to collect and process massive amounts of data. This data can now be explored and viewed as benefits to technological advancements, bringing limitless advantages to the government, education, engineering, healthcare, and other data-driven practices.

To be familiar with the concept of having Large Data "big" defining by both in dimension and meaning, particularly in education, allows for the exploration and prediction of learners behaviour across an enormous variety of contexts, acquisition degree, personal backgrounds, reflective progressions, academic intentions, environmental characteristics, personal potentials, skills –eve (Naga, I. and Hao, Y. 2014; Nessi. 2012).

The majority of these data considerations are currently being investigated in the education field to help devise instructional methods, analyze and measure the impact of these approaches on both learners and educators, and in general, all of which will help create a transformation in the educational sector by developing an effective learning environment.

3 PRESCHOOL EDUCATION ANALYTICS

In preschool, pupils are the main body of learning. Both teaching activities should be carried out to promote children's progress. A children-centered curriculum focuses on transforming passive education in children into constructive knowledge through the education process.

The unpredictability of learning activities puts forward high requirements for teachers. Big data analytics can offer a vast number of learning opportunities to all pupils, learning experiences can be designed in real-time based on the children's reactions to the learning activities (Ling Jin,2019).

When the number of states that have publicly subsidized preschool education increases, states are developing monitoring mechanisms to help assess program success and, as a result, how efficiently the public's money is used (By Shannon Riley-Ayers, Ellen Frede, W. Steven Barnett, and Kimberly Brenneman,2011).

In reality, the majority of state assessments of preschool systems are less than comprehensive in terms of science criteria, with many possessing such defects that understanding their findings is severely limited (Gilliam, W. S. & Zigler, E. F. 2000) (Gilliam, W. S. & Zigler, E. F. 2004).

Early childhood education systems are under - demand to gather data on both teachers and students and to use the data to make decisions.

Two significant impediments can prohibit early childhood educators from effectively using data to guide decisions. The first is a scarcity of literature on the best methods for data use in early childhood education. The second issue is a lack of capacity among preschool systems to collect data and use the findings to make decisions (Yazejian, N., & Bryant, D. 2013).

4 BIG DATA DIMENSIONS

In one view, Doug Laney (Doug Laney, 2011) describes Big Data as having three dimensions: volume, variety, and velocity. Thus, International Data Corporation (IDC) defined it: "Big data technologies refer to a new wave of technologies and architectures that allow high-velocity data collection, exploration, and, or analysis to economically extract value from enormous quantities of a wide variety of data." Two more characteristics seem to be significant: value and complexity. In the following paragraph, we will outline these characteristics.

<u>Three Vs</u>

Many examples of big data concentrate on the amount of data in storage, and size does matter, there are other essential characteristics of big data, such as data variety and velocity. The three Vs of big data (volume, variety, and velocity) (K. U. Jaseena and J. M. David,2014) constitute a thorough overview and dispel the misconception that big data is solely about data volume. In addition, each of the three Vs has its implications for analytics (See Figure 1).

Big data is everywhere over us, even though we don't recognize it right away. Part of the issue is that, even under exceptional circumstances, the rest of us do not agree with vast quantities of data in our everyday lives. We frequently struggle to grasp both the possibilities and threats posed by big data because we lack this immediate knowledge. As a result of these characteristics, there are currently many disagreements and difficulties in overcoming these characteristics in the future.

- ✓ **Data Volume:** That means it assesses the amount of data accessible to an organization and does not have to own any of it as long as it has access to it. The importance of diverse data records will decline as data volume grows concerning age, richness, form, and quantity, among other factors(A. TOLE,2013) (D. Laney,2001).
- ✓ **Data Velocity:** Data velocity refers to how quickly data is created, streamed, and aggregated. The speed and richness of data used by a variety of private transactions have steadily increased(for example, website clicks). Data velocity control is more than just a bandwidth problem; it's also an issue with data that's been swallowed (extract-transform-load) (K. S. Dr Jangala., M. Sravanthi, K. Preethi and M. Anusha,2015).
- ✓ **Data Variety:** Data variety is a measure of the richness of the information representation – text, images video, audio, etc. From an analytic perspective, it is possibly the principal obstacle to effectively using large volumes of data.

  Non-structured data, incompatible data arrangements, and ambiguous data semantics all pose major obstacles to analytic expansion (K. U. Jaseena and J. M. David,2014).
- ✓ **Data Value:** This stage assesses the usefulness of data in decision-making (K. U. Jaseena and J. M. David,2014). It has been noted that "the purpose of computing is insight, not numbers".

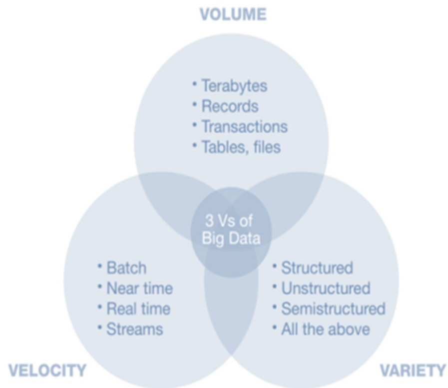

Figure 1: Three Vs of Big Data

  Data science is the study of and applying data to recognize it, but "analytic science" involves the statistical monitoring of large amounts of data.
- ✓ **Complexity:** The degree of interconnectedness is measured by complexity (probably very immense) and interdependence in large data structures such that a slight modification (or arrangement of small changes). Rather big changes or a minor transition that ripple through the system and significantly marks its behaviour, or no change at all, can be achieved by changing one or a few elements (Katal, A., Wazid, M., & Goudar, R. H. 2013).

5 BENEFITS OF BIG DATA IMPLEMENTATION IN PRESCHOOL

We will gain a thorough understanding of the educational process and its mechanism by using Big Data's different methods of data analysis.

The following are some of the most popular uses for Big Data in the preschool sector:

Collaboration: Many experts from various fields and backgrounds may share their knowledge and establish collaborative centres. These programs improve educational methods by encouraging teamwork and coordination, as well as coordinating academic views (Zweig, J., Irwin, C. W., Kook, J. F., & Cox, J. 2015).

Understanding the learning process: Any features, patterns, and learning speeds are unique to each learner. We will derive various recommendations to improve the learning process by providing insight into the educational learner's direction.

Feedback: The failure to recognize the causes of the issue and how to solve it is one of the most common challenges that the school system faces. We may

immediately determine the reason for a learner's failure using the conventional approach, but the learner can fail again. To address the issue, modern techniques propose extracting previously collected data from each learner's route and offering alternate methods of correction (Franco, Pablo Daniel, Antonio Matas, and Juan José Leiva,2020).
Predicting: Based on their profiles, specialists may create a viable curriculum performance for the next generation, using digital traces preserved from past interactions with the platform, as well as other evidence (Burchinal, M. R., Kainz, K., & Cai, Y. 2011).
Motivation: after seeing the implications of extensive data adoption, learners can see the gain and become more persuaded that they can see the impact of how this happens (Theodotou, E. 2014).

6 EDUCATIONAL BIG DATA NEEDS, OPPORTUNITIES, AND CHALLENGES

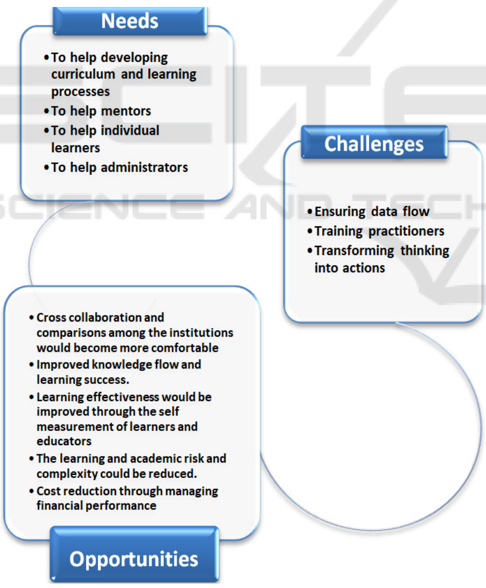

Figure 2: Educational Big Data Needs, Opportunities and Challenges adapted from (Shikha Anirban,2014)

For educators, big data analytics is now posing a significant challenge. People are currently concerned about institutions' intelligent outcomes of learning about pupils' learning and research institutions' blind pursuit of big data analytics without understanding the ramifications (Shikha Anirban,2014).

Figure 2 represents the educational (learning and academic) analytics from three aspects: Needs, Opportunities, and Challenges.

7 BIG DATA PROCESS AND CHALLENGES IN EDUCATION

7.1 Collection: Challenges and Propositions

The first step in revealing the importance of Big Data is to gather data. This necessitates the identification of data that may reveal useful and important knowledge. Data must be filtered for relevance and stored in a usable format since there is no point in spending money on massive amounts of data and computing infrastructure if the vast majority of the data in it is useless.
There are also difficulties with the accuracy of the data that has been gathered and described. Since the quality of data obtained by Big Data is entirely dependent on the quality of data gathered and the robustness of the procedures or metrics used, inter (national) comparison and assessment is complicated.
challenges

1. ***Ethics:*** Identifying the institute methods for protecting personal learner data protection, human approval, data ownership, and accountability are among the ethical issues for Big Data Analytics in Preschool. These issues exist when educational data collection and use are not subject to any formal ethics review. Moreover, difficulties arise when the data sources are complex and sophisticated (Willis, J.; Campbell, J.; and Pistilli, M. 2013)(Royal Statistical Society 2015).
2. ***Heterogeneity:*** The heterogeneity of data sources is the most important issue in the data collection process. (Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. 2014) Because of the diversity, representation, and semantics of the data sources, heterogeneous data complications rise. Furthermore, the majority of the new data generated is fundamentally different from the data forms on which the original structures were developed. Date formats and character fields are the most prevalent sources of representational errors. Database creators may attempt to link datasets using the student's name and surname to extract some vital information. Since character fields

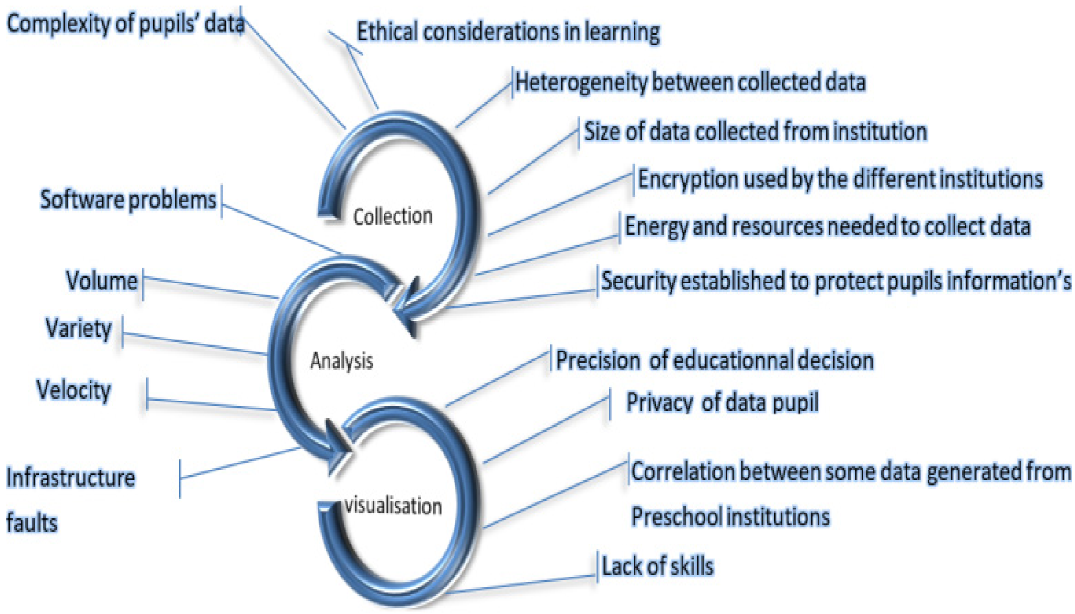

Figure 3: Big data process and challenges tree in education by authors

are item sensitive, even minor misrepresentation, such as using different capital letters, can slow down joins and searches.

3. ***Size of data:*** Another major issue is transferring the captured data. The speed at which the data is transferred could be a bottleneck in the process due to the size of the data (Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. 2014).
4. ***Energy and resources***: Storing and loading such a large amount of data necessitates a significant amount of energy and money. Finding the best-located servers to store the data is one of the challenges of Big Data. In addition, the server sites must be energy-efficient and flexible. The location is essential due to the speed of transfer of the stored data to do the analyses.
5. ***Encryption:*** Many secure transmissions necessitate some kind of encryption that must be agreed upon in advance by scholar institutions and learners' parents; however, institutions do not create commons laws to justify such measures and procedures(B.Tulasi,2013)
6. ***Data security:*** The unregulated accumulation of data by various organizations is perhaps the most severe threat to personal security (game application, social media…). This information raises serious security concerns, particularly because too many people voluntarily hand over that information (Luo Ying,2016). There are also questions about data fidelity and accuracy, as well as distribution, expiration, and access.

- **Propositions**

i. Several big data educational programs must be safeguarded in terms of privacy and security laws and procedures. IDC suggested five levels of increasing security: Privacy, compliance-driven, custodial, sensitive, and lockout are all words that come to mind when thinking about data security. Further research is needed to clearly identify these protection levels and compare them to existing law and analytics.

ii. For these reasons, ethical standards are required to ensure that data stewardship and ownership are established, and privacy concerns are addressed, ensuring that data is shielded from misuse. Educators and organizations will help educate curriculum design and pedagogy in this way, thus allowing students to become more aware of their learning habits.

iii. Data users and developers must be mindful that data must be changed and modified regularly to be used effectively.

iv. Since storing large amounts of data can be costly, some organizations have attempted to retain only a portion of the data gathered. As a result, interpreting data extraction findings may be more difficult; on the other hand, a portion of false connections and unexplained data ties can occur.

v. If the information is used for the benefit of the learners, such as predicting student actions and presenting a series of recommendations and services based on Learning Analytics, or if it is used for research to satisfy Learning Analytics intentions, the searchers must clarify how the information is used. Also, define the time frame during which learners' data is held, as well as a deletion procedure.

7.2 Analysis: Challenges and Propositions

Once data has been transformed into a usable format, it must be processed to generate actionable information. However, as the quality of information becomes more diverse, handling and interpreting a diverse data set is becoming more complex. To understand the information that is meant to be transported by these data, the analysis must involve referencing, integrating, and correlating disparate data sets.

As a result, the complexity of Big Data has been coined. But, how can we ensure that all data of a particular nature is precise and reliable? Or, to put it another way, incorrect results aren't the only issue with big data analytics. The speed at which the tests are completed is the main issue. This is one of the big data's three V's. Volume, Variety, and Velocity are the most common definitions for these Vs.

- **challenges**

a. ***Variety and Volume:*** During the data collection and incorporation process, the terms "volume" and "variety" were discussed.

b. ***Velocity:*** Large data velocities not only do it refer to the flow of data from sources to databases, but also refers to the flow of information from databases to the final analytics result. The speed at which data is extracted and analyzed is the most significant competitive advantage that an enterprise can achieve.

c. ***Infrastructure faults:*** Huge amounts of data that are critical to a company's success must be stored and analyzed, which necessitates a large and complex hardware system. More hardware structures would be needed as more and more complex data is processed. A hardware device can only be trusted for a limited amount of time. Intensive use and, in rare circumstances, production flaws would almost inevitably trigger a machine failure (Bala M. Balachandran, and Shivika Prasad,2017).

d. ***Software problems:*** Data loss isn't necessarily a hardware problem. Software may even fail, resulting in irreversible and potentially unsafe data loss. When a hard disk fails, there is usually another one to back it up, so data is not destroyed; but, when the software fails due to a programming "bug" or a design mistake, information is lost permanently. Hardware features restrict software solutions (CTI,10,2018).

- **Propositions**

i. Correct and timely decisions would maximize the return on investment and the institutions' educational-solution share. However, to examine the dynamic development of big data in education, a long-term financial plan is needed.

ii. To prevent such disasters, they use a backup mechanism that performs the primary task of saving all records. Companies gain continuity as a result of this, even though they are temporarily retracted.

iii. To address this problem, programmers devised a set of tools that would mitigate the consequences of a technological malfunction. Microsoft Word, for example, periodically saves the work that a user is performing to avoid data loss in the event of hardware or software failure. This is the fundamental concept behind avoiding total data loss.

7.3 Visualization: Challenges and Propositions

This is the final point, in which the evaluated data is made available to users in an interpretable format that can be implemented into existing systems and then used to direct decision-making.

We can resume the data visualization process by several steps noted below:

Prediction: Learner habits and potential success can be revealed by anticipation. As a result, appropriate action will help Learning Analytics achieve its objectives.

Intervention: Identify scholars who may be at risk, provide advice to learners who may need extra support, and help students succeed.

Recommendation: The ability to make recommendations to learners based on their activities is the primary objective of Learning Analytics.

Personalization: accelerate academic creativity and increase educational results. it can be done, for example, by personalizing e-learning based on a learner's skill or by assisting students with personalized learning tips.

Reflection and Iteration: Reflection aims to analyze previous work to enhance future experiences and to transform it into learning through personalization and adaptation. This iteration will maximize all Learning Analytics stakeholders in the design of its loop.

Benchmarking: Benchmarking is a method of learning that recognizes the best practices that yield superior outcomes. These activities are being repeated to improve success outcomes (Vorhies DW, Morgan NA. 2005).

- **challenges**

a. It is important to understand the degree of exactness that the handler needs to identify adequate data to produce an estimation or forecast of the specific likelihood and precision of a given outcome.
b. Determine what amount of time evidence is still operational and whether the relevance of publicly available findings expires.
c. To protect learner privacy, Big Data in Education requires disclosure, which makes the identity of the student available to disclose decisions. Furthermore, using Big Data Analytics implementations to forecast learners' potential educational results, realistic efficiency, and engagement could jeopardize their privacy.
d. Because of the strong connection, the prototype's removed details may have become improper statistical ones that exposed the erroneous correlation and misleading linkage to the specified component.
e. Big data poses challenges for data scientists and managers at any stage of the analytical process. Institutions face greater problems in hiring qualified data scientists who can work professionally with big data than in the analytical process itself (Daniel, B. 2014) (NARST 2015).

- **Propositions**

i. Create appropriate structures to efficiently manage the results.
ii. The establishment of a common standard based on internationally agreed-upon fundamental values would be a huge advantage as well. These revisions should include important existing considerations, circumstances, and complexities, as well as the development of specific protocols and the provision of acceptable conversion times for implementing the necessary modifications.
iii. States, especially high-tech firms, must have more financial capital to stimulate investment in this domain and to simplify laws by creating less complicated procedures.
iv. Sensitize parents and educators by assisting organizations in manipulating their knowledge and outcomes, as they play an important role in facilitating the first step for scientists
v. Institutions all over the world must reduce the number of methods and techniques used in the educational arena and work together to develop international rules that can standardize the complex data process and increase the efficacy of the outcomes.

***Common challenges*:** *Build a specialized centre over the world to overcome the lack of Big Data researchers in education, and promote the competence of the staff during each process of treatment.*

8 CONCLUSION

While we can establish and incorporate effective solutions to address learning barriers from preschool to high school by concentrating efforts on the child learning process, we discovered a lack of studies that shed light on big data analytics in preschool (BDTP) problems process including their challenges.

This paper kicks off a joint exploration campaign to look at (BDTP) questions and problems. Several significant problems in big data collection, administration, and implementation have been identified. That is something we think needs to be tackled over the next decade. Our upcoming study will concentrate on gaining a skilled understanding of the problems associated with (BDTP). Via our research, we will continue to investigate more effective solutions to some of the issues posed in this article.

Table 1: Summarizing the Grid of big data analytics challenges and propositions by authors

	Process issues	Challenge	propositions	Comments (+): strongness (-): weakness
Process of Big data Analytics in preschool with some propositions	**collection**	***Ethics***	enabling the pupil to become further conscious of their own learning behaviors	(+) give learners a self-confident and protect their rights (-) require collaboration
		Heterogeneity	Establish a periodic revise of DATA stored	(+) simple to update (-) application require a periodic modification
		Size of data	Storing a part of the data collected	(+) resolve a part of whole problems (-) losing some sensible information
		Energy and resources:	Looking for an approximate location to minimize distance between stations	(+) accelerate the process of data acquisition (-) very costly
		Encryption:	Establish an efficient common encryption strategy	(+) Has huge potential, can protect data versus attackers (-) Has limitations outcomes
		Data security	Increasing security levels and standardizes it	(+) efficient mechanism to protect data (-) more sustained efforts towards implementation are required
	analytics	***Variety and Volume***	Follow and examine data process progression	(+) very important to start any analytics process (-) lack of indicators precision
		Velocity	Corporation between organization to plan the fast decisions	(+) reach certified results (-) require a cost resources
		Infrastructure faults	Develop a series of tools to reduce the impact of hardware failure	(+) provide a useful mechanism to gain additional time (-) insufficient budget
		Software problem	use a backup system	(+) reduce the rate of failure (-) limited update
	visualization	***level of precision***	Design appropriate systems and minimize strategies around the word	(+) reach an operational results (-) less level of collaboration
		Data expiration	Allow sufficient transition periods to apply the convenient changes	(+) keep the suitable and useful data (-) Not feasible by the majority of institution
		Privacy	Sensitize the parents and learners and demand their permission	(+) protect pupil's confidentialities (-) need collaboration
		investment	provide more financial resources	(+) the essential pillar for each changes (-) rare attractive results
Common challenge: we can consider lack of skilled big data researchers in education. The most shared factor between all the processes of treatment In big data analytics.				

REFERENCES

A. Franco, Pablo Daniel, Antonio Matas, and Juan José Leiva. 2020. *" Big Data irruption in education". Pixel-Bit. Revista de Medios y Educación 57: 59–90.*

Reidenberg, Joel R., and Florian Schaub. 2018. *"Achieving big data privacy in education". Theory and Research in Education 16: 263–79*

Zweig Clare W. Irwin Janna Fuccillo Kook Josh Cox, *"Data collection and use in early childhood education programs": Evidence from the Northeast Region Jacqueline Education Development Center, Inc. In collaboration with the Early Childhood Education Research Alliance*

Naga, I. and Hao, Y. (2014). *" A Holistic Analysis of Cloud-Based Big Data Mining". International Journal of Knowledge, Innovation, and Entrepreneurship. Volume 2 No. 2, 2014, pp. 56—64. Retrieved from: http://ijkie.org/IJKIE_December2014_IRINA&HAO.pdf*

NESSI. (2012). *Big Data: A New World of Opportunities. Retrieved from:http://www.nessi-europe.com /Files/Private/NESSI_WhitePaper_BigData.pdf.*

Ling Jin, *"Investigation on Potential Application of Artificial Intelligence in Preschool Children's Education", To cite this article: 2019 J. Phys.: Conf. Ser. 1288 012072*

By Shannon Riley-Ayers, Ellen Frede, W. Steven Barnett, and Kimberly Brenneman, *"Improving Early Education Programs through Data-based Decision Makin" (2011).*

Gilliam, W. S. & Zigler, E. F. (2000). *A critical meta-analysis of all evaluations of state-funded preschool from 1977-1998: Implications for policy, service delivery, and program evaluation. Early Childhood Research Quarterly, 15(4), 441-473.*

Gilliam, W. S. & Zigler, E. F. (2004). *"State efforts to evaluate the effects of prekindergarten": 1977-2003. Retrieved September 10, 2008.*

Yazejian, N., & Bryant, D. (2013). *Embedded, collaborative, comprehensive: One model of data utilization. Early Education and Development, 24(1), 68–70.*

Doug Laney *"Pattern-Based Strategy: Getting Value from Big Data.," Gartner Group press release. . Available at http://www.gartner.com/it/page.jsp?id=1731916, July 2011*

K. U. Jaseena and J. M. David, *"ISSUES, CHALLENGES, AND SOLUTIONS: BIG DATA MINING," Computer Science & Information Technology (CS & IT), no. pp. 131–140, 2014*

A. TOLE, *"Big Data Challenges," Database Systems Journal, vol. Vol. IV, no. no. 3, pp. 31-40, 2013*

D. Laney, *"3D Data Management: Controlling Data Volume, Velocity, and Variety," Stamford, February 6, 2001.*

K. S. Dr. Jangala., M. Sravanthi, K. Preethi and M. Anusha, *"Recent Issues and Challenges on Big Data in Cloud computing," IJCST, vol. Vol. 6, no. Issue 2, April - June 2015*

K. U. Jaseena and J. M. David, "ISSUES, CHALLENGES, AND SOLUTIONS: BIG DATA MINING," Computer Science & Information Technology (CS & IT), no. pp. 131–140, 2014.

Katal, A., Wazid, M., & Goudar, R. H. (2013). *"Big Data: Issues, Challenges, Tools and Good Practices". IEEE, 404-409.*

Zweig, J., Irwin, C. W., Kook, J. F., & Cox, J. (2015). *"Data collection and use in early childhood education programs": Evidence from the Northeast Region (REL 2015–084). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.*

Franco, Pablo Daniel, Antonio Matas, and Juan José Leiva. 2020. *Big Data irruption in education. Pixel-Bit. Revista de Medios y Educación 57: 59–90*

Burchinal, M. R., Kainz, K., & Cai, Y. (2011). *How well are our measures of quality predicting to child outcomes: A meta-analysis and coordinated analyses of data from large scale studies of early childhood settings. In M. Zaslow, I. Martinez-Beck, K. Tout & T. Halle (Eds.), Measuring quality in early childhood settings (pp. 11–31). Baltimore: Brookes Publishing.*

Theodotou, E. (2014). *"Early Years Education: are Young Students Intrinsically or Extrinsically Motivated Towards School Activities? A Discussion about The Effects of Rewards on Young Children's Learning". Research in Teacher Education.Vol: 4 No:1 pp: 17–21.*

Shikha Anirban, *"Big Data Analytics in the Education Sector: Needs,Opportunities and Challenges", International Journal of Research in Computer and Communication Technology, Vol 3, Issue 11, November –2014*

Willis, J.; Campbell, J.; and Pistilli, M. (2013). *ethics, big data, and analytics: an application model. EDUCAUSE. Retrieved July 11, 2017, from http://er.educause.edu/articles/2013/5/ethics-big-data-and-analytics-a-model-for-application.*

Royal Statistical Society (2015)."*The opportunities and ethics of big data. "report. Retrieved July 11, 2017, from http://www.rss.org.uk/Images/PDF/ influencing-change/2016/RSS-report-apps-and-ethics-of-big-data-Feb-2016.pdf.).*

Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). *"Big data and its technical challenges". Communications of the ACM, 57(7), 86–94. doi:10.1145/2611567*

B.Tulasi, *"Significance of Big Data and Analytics in Higher Education" , Journal of Computer Applications (0975 – 8887) Volume 68– No.14, April 2013 21.*

Luo Ying. *"Research on Big Data Security and Privacy Protection (J). Information Communication", 2016(1):162-163.*

Bala M. Balachandran, and Shivika Prasad, *" Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence", International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, KES2017, 6-8 September 2017, Marseille, France.*

CTI,10 *Common Causes of Data Loss: https://consoltech.com/blog/10-common-causes-of-data-loss/*

Vorhies DW, Morgan NA. 2005. *Benchmarking marketing capabilities for sustained competitive advantage. Journal of Marketing 69(1): 80 – 94*

Daniel, B. (2014). *"Big data and analytics in higher education: opportunities and challenges". British Journal of Educational Technology, 46(5), 904-920.*

NARST (2015). *"Big data and learning analytics: a new frontier in science and engineering education research". Symposium presented at the annual meeting of the National Association for Research in Science Teaching. Chicago, IL 1-11, Retrieved June 6, 2017, from: https://concord.org/sites/default/files/pdf/big-data-and-learning-analytics.pdf.*