Highlighting the Challenges of Big Data Analytics Implementation in Mainstream Industries

Mary Jane C. Samonte
School of Information Technology,
Mapúa University, Manila, Philippines
mjcsamonte@yahoo.com

Johndale Alfred A. Julian
School of Management and Information Technology, De La Salle-College of Saint Benilde
School of Information Technology, Mapúa University, Manila, Philippines
johndalealfred.julian@benilde.edu.ph

Abstract

In the past few years, Big Data Analytics has advanced in improving firms' decision-making capabilities and developing and optimizing the day-to-day operations of other mainstream industries. The extent of implementing Big Data and Big Data analytics has become unimaginable in comparison to what industries use it for in its early stages. This paper highlights the increase of Big Data Analytics implementation throughout several industries and practices. This paper aims to highlight the innovation and improvements Big Data Analytics has brought upon these industries, as well as the challenges of implementing Big Data Analytics.

Keywords and Phrases
Big Data, Internet-of-Things, Smart Cities, Decision-making capabilities, Analytics.

1 Introduction
According to Chris Anderson's *Wired Magazine* editorial from 2008, we now live in the 'Petabyte Age.' (Anderson 2008). The adjective 'big' does not even come close to adequately describing the amount of data that is being collected, processed, loaded, and analyzed nowadays. The Large Hadron Collider located in CERN alone had up to 15 petabytes of data in 2010, while CISCO obtained over 667 exabytes of data in 2013. The idea and concept of Big Data Analytics implementation already existed even before we knew it. However, the term has since then become famous as technology exponentially evolved throughout the years.

When talking about hundreds of petabytes of data, it is not just a simple data set that we can easily collect, ingest, and analyze as we did with traditional database systems (Kaisler et al. 2013). Big Data calls for a revolutionary approach that data specialists have worked on for the past few years to catch up with the ever-growing demand for data analysis. The continuous evolution of technology has undoubtedly brought the invaluable importance of relying on data to improve decision-making capabilities and develop a plethora of automated and intelligent systems for our disposal. The role of big data has never been more significant in today's era ranging from a system recommending what one should watch next on Netflix to automating systems in a developing Smart City.

According to Google Trends, the rise of interest in big data worldwide began growing in the early 2011s and has consistently grown since. Statista says the global big data analytics market will continue to grow at a 30% rate, with revenue coming close to 68 billion USD by 2025 (Statista, n.d.). This shows that big data analytics has become more prominent and that more and more companies and industries are continually jumping into the hype of implementing big data analytics.

Some of the biggest industries that invested early on in big data analytics implementation are healthcare (Viceconti et al. 2015), agriculture (Kamilaris et al. 2017), military (Song et al. 2015), banking (Hassani et al. 2018), among several others. These industries may have different specific purposes for utilizing big data analytics implementation.

However, all are towards developing and improving their traditional operations, solving issues, and answering questions that only big data analytics can do.
Studies that focus on implementing big data in industries were collected and analyzed to provide a general idea of the challenges several industries face when it comes to implementing big data analytics.

1.1 Objectives
With the growing demand of organizations to implement Big Data capabilities, organizations who are just starting to ride into the hype of Big Data may encounter new challenges that could prove to be counter-productive that may defeat the purpose of implementing such technologies.

This study aims to conduct the following: 1) Determine major industries that aims to implement or already implements Big Data; 2) Determine the common types of challenges that Big Data implementers have experienced, or are currently experiencing, and; 3) Propose a model to prevent and/or mitigate the challenges of Big Data operations and implementations.

2 Literature Review
This section of the study aims to determine the fundamental reasons for implementing big data.

2.1 The need for Big Data
There are a several fundamental reasons on why organizations to implement big data capabilities. The ever-changing business models of every industry have been vastly expanding which introduces new opportunities. But with these improvements and enhancements has its own corresponding challenge that comes along with it (Kezunovic et al. 2020). The impact of big data analytics has proven to be quite the transformational stride in every industry that seek innovation and better decision-making models for their operations in the long run.

2.2 Challenges of implementing Big Data
The goal of Big Data Analytics is to draw out value from data. With data being proven to be a sensitive and complex domain, the Data Management Lifecycle was introduced so that proper handling and management of data would take place in order to make the most value out of it while keeping the usage and its use-cases within its allowable boundaries.

Due to the fact that data domain can be applied to a vast array of industries and operations, complications and problems often arise due to the complexity of such technologies when being implemented. This is not limited to just the technical side of things, since every aspect of a business operation could be translated into data and that could result into complexities due to a several factors (Kezunovic et al. 2020).

3 Methods
The collected related studies used for this paper were obtained from reputable publications such as Elsevier, ScienceDirect, Journal of Computational Science, and Image and Vision Computing. The initial steps of collecting studies were to search databases and search for the exact term "big data." The search for related studies was filtered to limit only those studies published from 2017 to 2020. Exploratory data analysis using Python and various data science libraries was further conducted to extract the major challenges that most industries are experiencing.

4 Data Collection
Among the collected twenty-two (22) studies, seventeen (17) papers qualified to be under the criteria of application of Big Data or Big Data Analytics in a specific industry. A total of fifteen (15) studies were retrieved from Elsevier and two (2) from ScienceDirect. As for filtering papers, their title or abstract must focus on big data implementation in a specific industry.

For example, the study of Dai et al., 2020 is titled *"Risk assessment of agricultural supermarket supply chain in big data environment."* Some studies that did not focus on applying big data analytics in a particular industry but can still be deemed qualified under the criteria due to its widespread application in a general sense. The qualified studies were then categorized by the identified industry where big data analytics were implemented or focused, as seen in Table 1.

Table 1. Collected studies categorized by the industry where big data analytics is applied

Title of Study	Publisher	Industry focused
Atitallah et al. (2020) Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions	Elsevier	Smart Cities Development
Barham and Daim (2020) The use of readiness assessment for big data project	Elsevier	Smart Cities Development
Dai and Liu (2020) Risk assessment of agricultural supermarket supply chain in big data environment	Elsevier	Agriculture
Gbadamosi et al. (2020) Big data for Design Option Repository: Towards a DFMA approach for offsite construction	Elsevier	Architecture
Guo et al. (2020) Research on an advanced intelligence implementation system for engineering process in industrial field under big data	Elsevier	Engineering/Construction
Haseeb et al. (2020) EBDS: An energy-efficient big data- based secure framework using Internet of Things for green environment	Elsevier	Natural environment
Kezunovic et al. (2020) Big data analytics for future electricity grids	Elsevier	Electricity
Kong et al. (2020) A systematic review of big data-based urban sustainability research: State-of-the-science and future directions	Elsevier	Urban development/planning
Pérez-Chacón et al. (2020) Big data time series forecasting based on pattern sequence similarity and its application to the electricity demand	ScienceDirect	Electricity
Sellami et al. (2020) On the use of big data frameworks for big service composition	Elsevier	General/Big Service
Shamim et al. (2020) Big data analytics capability and decision-making performance in emerging market firms: The role of contractual and relational governance mechanisms	Elsevier	General
Song et al. (2020) Intelligent state of health estimation for lithium-ion battery pack based on big data analysis	Elsevier	Automotive
Suoniemi et al. (2020) Big data and firm performance: The roles of market-directed capabilities and business strategy	Elsevier	General
Tang et al. (2020) Trends Prediction of Big Data: A Case Study based on Fusion Data	Elsevier	Technology
Wang et al. (2020) A Hybrid Big Data Analytical approach for Analyzing Customer Patterns through an Integrated Supply Chain Network	ScienceDirect	Supply Chain
Yang et al. (2020) Impacts on environmental quality and required environmental regulation adjustments: A perspective of directed technical change driven by big data	Elsevier	Natural Environment
Zhu and Jiang (2020) Optimization of face recognition algorithm based on deep learning multi feature fusion driven by big data	Elsevier	Technology

5. Results and Discussions

This study aims to highlight the challenges of mainstream industries in implementing or managing big data analytics implementation. With the seventeen (17) selected published studies, those discussing their encountered challenges were further examined. The challenges encountered by specific mainstream industries were then categorized into major topics such as technological, management, business-centric, cost, among several others.

5.1 Numerical Results

5.1.1 Categorizing the Challenges Encountered per Industry

According to the study of Atitallah et al. (2020), they encountered several challenges in big data analytics implementation in developing a smart city ecosystem. They categorized it into two major factors: Technical and Business Challenges.

The selected studies were further examined to determine the significant challenges encountered and highlight the general categories encountered per industry, as seen in Table 2 and categorized in Table 3.

Table 2. Challenges encountered per industry according to study

Study	Industry focused	Challenges Encountered
Atitallah et al. (2020)	Smart City Development	Technical, Business
Dai and Liu (2020)	Supply Chain	Technical, Operational
Gbadamosi et al. (2020)	Architecture	Technical
Guo et al. (2020)	Engineering	Technical
Haseeb et al. (2020)	Natural Environment	Physical, Security
Kezunovic et al. (2020)	Electricity	Physical, Operational, Security
Kong et al. (2020)	Urban Development/Planning	Physical, Technical, Security
Pérez-Chacón et al. (2020)	Electricity	Technical
Shamim et al. (2020)	General	Management
Song et al. (2020)	Automotive	Technical, Physical
Suoniemi et al. (2020)	General	Technical, Physical, Operational, Business
Wang et al. (2020)	Supply Chain	Technical, Security
Yang et al. (2020)	Natural Environment	Natural
Zhu and Jiang (2020)	Technology	Technical

Table 3. Major categories of challenges encountered

Challenges	Industries Affected	Studies covered
Technical (7*)	Smart City Development Supply chain Architecture Engineering Electricity Automotive Supply Chain Technology	(Atitallah et al. 2020), (Dai and Liu 2020), (Gbadamosi et al. 2020), (Guo et al. 2020), (Pérez-Chacón et al. 2020), (Song et al. 2015), (Suoniemi et al. 2020), (Wang et al. 2020), (Zhu and Jiang 2020), (Barham and Daim 2020)
Business/Management (1*)	Smart City Development	(Atitallah et al. 2020)
Security (4*)	Natural Environment Electricity Urban Development/Planning Supply Chain	(Haseeb et al. 2020), (Kezunovic et al. 2020), (Kong et al. 2020), (Wang et al. 2020)
Operational (2*)	Supply Chain Electricity	(Dai and Liu 2020), (Kezunovic et al. 2020)
* - number of industries affected		

Technical challenges, as shown in Table 4. have been the most dominant type of challenge being experienced by most industries in implementing Big Data Analytics. Given that Big Data Analytics implementation is very technical at a large-scale, it is expected that technical issues would be quite the common issue being encountered by big data implementations by different industries shown in Figure 1. Operational challenges would encompass business and physical challenges in this case as it covers most of the smaller-scale issues at hand including but not limited to costs, quality-of-service, and planning concerns. Security challenges come second after technical problems being encountered by most industries implementing Big Data, as data lifecycle proves to be quite a strenuous domain to handle, not to mention common security issues needed for infrastructures that run big data analytics.

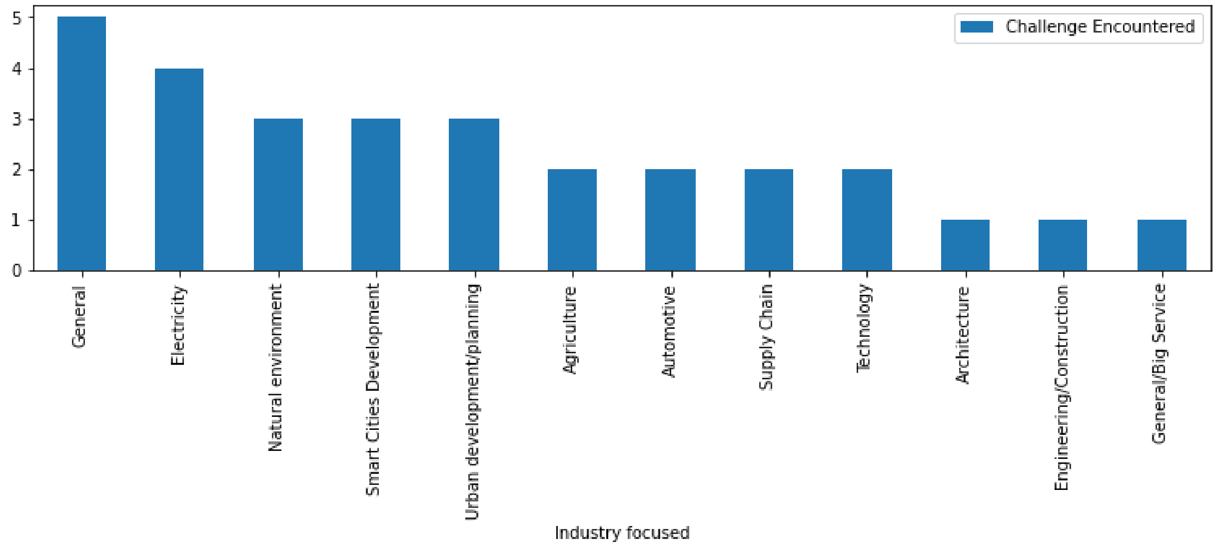

Figure 1. Industries with different challenges encountered in implementing Big Data Analytics

Table 4. Count of challenges encountered per Industry, and most type of challenge encountered

Industry focused	Count	Most Type of Challenge Encountered
Agriculture	2	Technical
Architecture	1	Technical
Automotive	2	Technical
Electricity	4	Physical
Engineering/Construction	1	Technical
General	5	Management
Natural environment	3	Physical
Smart Cities Development	2	Technical
Supply Chain	2	Technical
Technology	1	Technical
Urban development/planning	3	Physical

5.2 Technical challenges

As seen in the previous table, technical issues are the most common challenges encountered by industries implementing Big Data Analytics, which is no surprise since Big Data is quite a technical realm after all. Implementing a sophisticated analytics system constitutes a massive volume of generated data being ingested, collected, and processed (Atitallah et al. 2020). The industries involved in tackling technical big data challenges are Smart City Development, Supply Chain, Architecture, Engineering, Electricity, Automotive, Supply Chain, and Technology industries.

5.2.1 Huge volume and mixture of Unstructured and Structured Data

In the study by Atitallah et al. (2020), one of the most technical issues they encountered were IoT big data issues such as communication, database management, and the various mixtures of structured and unstructured data (Atitallah el al. 2020). This particular issue is not uncommon to big data analytics in general, as this falls into the six V's of big data (Atitallah et al. 2020):

- **Variety**: the plethora of data formats categorized into unstructured, semi-structured, quasi-structured, and structured formats.
- **Variability**: the undeterminable shifting of data movement
- **Volume**: the massive amount of data being generated, collected, processed, and analyzed
- **Velocity**: how fast the data is being created and captured
- **Veracity**: the correctness and validity of data
- **Value**: what an organization can make out of the collected and processed data

The approach to this issue varies from one industry to another, as they have different situations regarding the six V's of their big data. However, the demands of big data implementation have required technology to come up with newer and modern software implementations and hardware platforms (Chebbi et al. 2015).

5.2.2Technology Limitations

From the several studies reviewed, a lot of them encountered challenges concerning technology limitations. Some industries attempt to implement quite advanced technologies such as artificial intelligence, machine learning, deep learning, neural networks, image recognition, speech recognition, among several others (Atitallah et al. 2020), (Dai and Liu 2020), (Gbadamosi et al. 2020), (Guo et al. 2020), (Perez-Chacón et all. 2020), (Song et al. 2015), (Suoniemi et al. 2020), (Wang et al. 2020), (Zhu and Jiang 2020), (Haseeb et al. 2020), (Kezunovic et al. 2020), (Kong et al. 2020).

These technologies are usually implemented when traditional analytics techniques are not viable due to the enormous volume and variety of data (Kezunovic et al. 2020). For these technologies to be valuable in big data analytics implementation, they require a higher level of precision and accuracy in predicting or succeeding in a specific operation.

5.2.3 Engineering Concerns

While big data analytics revolve around data science fundamentals, several engineering aspects still need to be implemented to be fully functional and valuable (Kezunovic et al. 2020). For example, data capture in specific industries requires highly specialized and sophisticated hardware such as sensors, scanners, radio frequency identifiers, and recognition systems.

One might say that the actual data science (which usually includes complex analysis, advanced statistics, and machine learning) is only 10% of the job, while the vast 90% remaining constitutes setting up the data pipelines, workflow, storage, computing, and processing (Kezunovic et al. 2020). Data scientists and engineers must also be familiar with the industry's niche. Each industry and practice require a certain level of understanding and knowledge with a certain level of information granularity from the scientists and engineers working around with its big data.

5.3 Security Challenges

One critical requirement that is often overlooked in big data analytics implementation is security (Haseeb et al. 2020). Security brings crucial importance linked to the data and connected data pipelines and systems (Kezunovic et al. 2020). The industries affected by security challenges in big data analytics implementation are the natural environment, electricity, urban planning/development, and supply chain. In big data implementations, security executes a critical function to mitigate potential threats and ensure that data is tamper-free, valid, and accessed only in authority. The absence or inadequacy of security in the entirety of big data implementation increases the risk of data leakage.

5.3.1 Cybersecurity

Distributed systems are more often than not employed in big data analytics as it utilizes an array of parallel hardware and highly complex software for a faster and more efficient big data storage and computing. Cybersecurity is a crucial component of distributed systems as to which the tools and systems needed should be vigorous to be able to resist unauthorized data access or even fabrication or tampering of data (Li et al. 2017).

5.3.2 Data Privacy

Today, data privacy is an ever-growing issue often associated with big data (Le Ray and Pinson 2020). If unauthorized data access has occurred, one could have malicious intent about specific organizational assets or consumer data (Le Ray and Pinson 2020). Systems that employ smart meters have been immensely helpful in ensuring data privacy in modern big data analytics systems. Cryptography algorithms that encrypt data into unintelligible languages also assist with the confidentiality and integrity of data from its collecting phase to its end phase (Haseeb et al. 2020).

5.4 Operational Challenges

The industries that had difficulties implementing big data analytics are the supply chain and electricity industries. While this is seemingly a non-technical challenge, integrating big data analytics into certain situations might differ in several industries based on their day-to-day operations, processes, and existing workflows. For instance, in the supply chain industry, big data analytics processes might vary depending on external factors such as natural, economic, political, and legal environment, which one might refer to as *Supply Chain Risk* (Dai and Liu 2020). There are also risks linked to production issues, such as data obtained could be unexpectedly affected by other uncontrollable events such as natural disasters (Dai and Liu 2020).

5.4.1 Distribution and Logistics

Distribution issues are also not an uncommon operational challenge that big data implementations go through in specific industries. Logistics processes play a huge role that makes implementing big data quite tricky as there are so many indeterminate variables such as longer distribution links and the considerable storage requirement for historical data, which might also lead to the risk of loss or tampering of data (Dai and Liu 2020).

5.4.2 Incompleteness of data

Since the supply chain organizations operate widely and often employ subgroups outside of their organization, managing data collection and ensuring its correctness can be a significant challenge to big data implementations. Loss of consistency in big data operations and processes might occur at the retail level or lower (Dai and Liu 2020).

5.5 Business Challenges

While this was just stated in one (1) study from the collected papers, we believe this is a big data implementation challenge that is often overlooked. Business challenges might vary depending on the organization and its management that is implementing big data capabilities. The study of Atitallah et al. (2020), highlighted issues under business challenges.

5.5.1 Cost

Costing might be the most common business challenge that big data implementations encounter. Big data analytics require sophisticated devices such as sensors, software, actuators, and specialists such as data scientists, data engineers, software engineers, among several others. These are the most crucial and significant factors that define the extensive costs of big data implementation (Atitallah et al. 2020).

5.5.2 Quality of Service

Establishing big data capabilities requires an extensive quality of service requirement that must be met at all times. QoS is a critical factor that employs a measurement system to validate the data quality, correctness, response time of processes and operations, scalability, and overall stability and reliability (Atitallah et al. 2020). Integration of technologies and operations must be considered to have accurate metrics that define good quality of service.

5.5.3 Planning

Planning requires the utmost importance in establishing big data analytics capability. The planning phase of any project bootstraps the whole operation moving forward and can be the deciding factor if a project is feasible or not and defines further success and its future potential. In smart cities development, each area must develop an execution and development plan that is accustomed to their citizens' needs (Atitallah et al. 2020).

6. Proposed Decision-Making Evaluation System for Big Data Implementation

It would be challenging to develop a blanket system that covers all specifications and needs of every industry implementing big data analytics. After all, every industry has different requirements, technicalities, challenges, purposes, and goals for developing big data analytics capabilities even at the level of the lowest unit in the hierarchy. For an organization to anticipate and overcome challenges with their big data analytics implementation, this study proposes a modified version of *Hierarchical Decision Model (HDM)* (Barham 2020*)* in conjunction with the proposed guidelines of the study of Okello-Orlale in 2018. The HDM method introduced by Barham determined essential steps that further enhances the judgement of organizations whether it is viable to implement Big Data Analytics while Okello-Orlale's study introduces several essential steps and guidelines in maximizing the potential of data driven policymaking. This also introduces the channel for the specialists' and experts' judgments to actively determine potential challenges and develop a set of decisions that could dynamically overcome the determined challenges (Okello-Orlale 2018). The responsibility framework and evidence base building guidelines introduced by Okello-Orlale is integrated with the previous Hierarchical Decision Model introduced by Barham 2020 as seen on Figure 2.

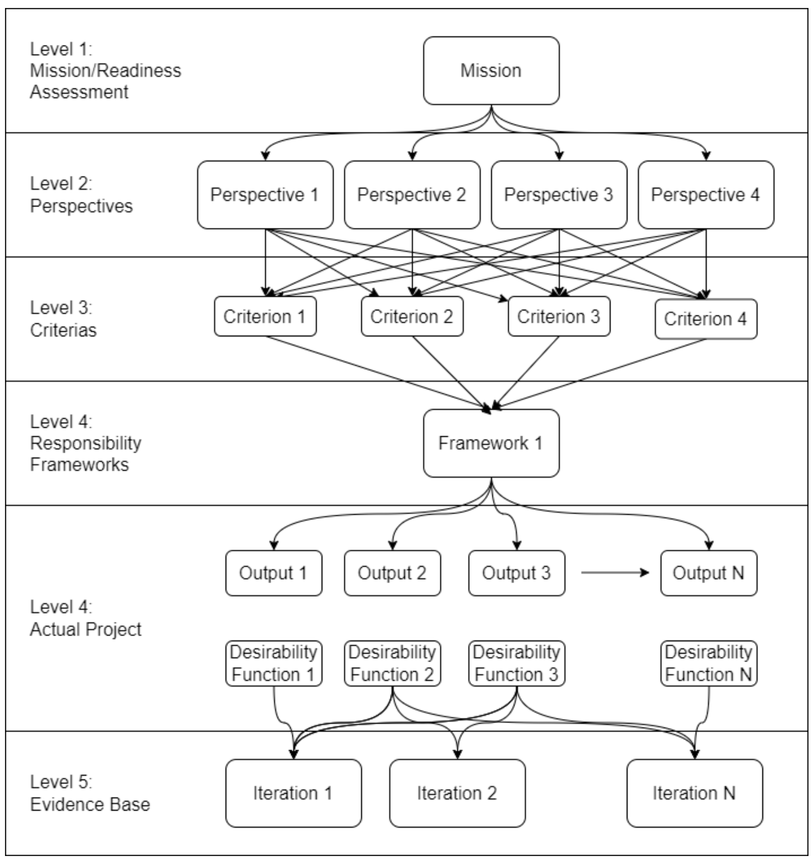

Figure 2. Hybrid Hierarchical Decision Model derived from Barham 2020 and Okello-Orlale 2018

Responsibility Frameworks proved to be an essential step in data-driven decision-making steps as it employs a standard in all aspects of implementation and operation of Big Data Analytics in organizations. Evidence Base was also integrated for future iterations of Big Data implementations as it records what is currently working with the current implementation and can be used in further updates and improvements.

The template should be managed and analyzed by industry-specific experts as they may have the depth of understanding and knowledge required to develop and manage a big data project in their organization/industry. Each industry will require a different input for each level in the hierarchy, thus requiring experts to carefully analyze each step moving forward to overcome the challenges they might encounter or currently encountering. Several studies have already implemented the model above in a specific industry, such as smart cities development (Barham and Daim 2020) and have successfully employed the HDM to build a readiness assessment system for data analytics implementation.

7 Implications and Conclusions

This study aimed to define the significant categories of challenges that industries face when employing big data implementations to improve and innovate their operations. With the four (4) defined key challenges described as Technical, Business, Operational, and Security, the study aimed to have a generalized yet systematic decision-making process to overcome these challenges.

While the model proposed above will require a significant amount of research and fine-tuning per specific industry or organization, the study proposes that sophisticated decision-making processes will help industry practitioners develop innovative and effective solutions.

8 Limitations and Future Research

The limitations of this research would be the small-scale number of research papers collected, which might have overlooked other considerable challenges those other industries are experiencing. Another method of searching could be employed, particularly looking up major industries and other categories of challenges being encountered. The model proposed above is also a generalized way of creating a systematic decision-making process. Future research could investigate fine-tuning the model into which industries are being affected per challenge category. This paper recommends creating a repository of derived models from the mentioned decision-making system above and creating a systematic analysis.

References

Anderson, C., The End of Theory: The Data Deluge Makes Scientific Method Obsolete, WIRED, Available: https://www.wired.com/2008/06/pb-theory/.

Atitallah, S. Driss, M. Boulila, W. and Ghezala, H., Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions, *Computer Science Review*, vol. 38, 2020.

Barham, H. and Daim, T., The use of readiness assessment for big data projects, *Sustainable Cities and Society*, vol. 60, 2020.

Big data analytics market revenue 2025, Statista, Available: https://www.statista.com/statistics/947745/worldwide-total-data-market-revenue/.

Chebbi, I. Boulila, W. and Farah, I., Big Data: Concepts, Challenges and Applications, *Computational Collective Intelligence, 7th International Conference, ICCCI 2015*, pp. 638-647, Madrid, Spain, September 21 – 23, 2015.

Dai, M. and Liu, L., Risk assessment of agricultural supermarket supply chain in big data environment, *Sustainable Computing: Informatics and Systems*, vol. 28, 2020.

Gbadamosi, A. Oyedele, L. Mahamadu, A. Kusimo, H. Bilal, M. Delgado, J. and Yakubu-Muhammed, N., Big data for Design Options Repository: Towards a DFMA approach for offsite construction, *Automation in Construction*, vol. 120, 2020.

Guo, Y., Zhang, B., Yu, S., and Kai, W., Research on an advanced intelligence implementation system for engineering process in industrial field under big data, *Expert Systems with Applications*, vol. 161, 2020.

Guo, Y., Zhang, B., Yu, S., and Kai, W., Research on an advanced intelligence implementation system for engineering process in industrial field under big data, *Expert Systems with Applications*, vol. 161, 2020.

Haseeb, K. Lee, S. and Jeon, G., EBDS: An energy-efficient big data-based secure framework using Internet of Things for green environment, *Environmental Technology & Innovation*, vol. 20, 2020.

Hassani, H. Huang, X. and Silva, E., Digitalisation and Big Data Mining in Banking, *Big Data and Cognitive Computing*, vol. 2, no. 3, 2018.

HimaBindu, G. Kumar, Ch. R. K. Hemanand, Ch. and Krishna, N. R., Hybrid clustering algorithm to process big data using firefly optimization mechanism, *Materials Today: Proceedings*, 2020.

Kaisler, S. Armour, F. Espinosa, J. A. and Money, W., Big Data: Issues and Challenges Moving Forward, *46th Hawaii International Conference on System Science*, Wailea, HI, USA, January 7-10, 2013.

Kamilaris, A. Kartakoullis, A. and Prenafeta-Boldu, F., A review on the practice of big data analysis in agriculture, *Computer and Electronics in Agriculture*, vol. 143, pp. 23-37, 2017.

Kezunovic, M. Pinson, P. Obradovic, Z. Grijavla, S. Hong, T. and Bessa, R., Big data analytics for future electricity grids, *Electric Power Systems Research*, vol. 189, 2020.

Kong, L. Liu, Z. and Wu, J., A systematic review of big data-based urban sustainability research: State-of-the-science and future directions, *Journal of Cleaner Productions*, vol. 273, 2020.

Le Ray, G. and Pinson, P., The ethical smart grid: Enabling a fruitful and long-lasting relationship between utilities and customers, IDEAS, Available: https://ideas.repec.org/a/eee/enepol/v140y2020ics0301421520300185.html, Accessed on October 21, 2020.

Lian, Y. Zhang, G. Lee, J. and Huang, H., Review on big data applications in safety research of intelligent transportation systems and connected/automated vehicles, *Accident Analysis & Prevention*, vol. 146, 2020.

Li, Z. Shahidehpour, M. and Aminifar, F., Cybersecurity in Distributed Power Systems, *Proceedings of the IEEE*, vol. 105, no. 7, pp. 1367-1388, 2017.

Narayanan, U. Paul, V. and Joseph, S., A novel system architecture for secure authentication and data sharing in cloud enabled Big Data Environment, *Journal of King Saud University – Computer and Information Sciences*, 2020.

Okello-Orlale, R., Unpacking the disruptive potential of data driven policy making process and blockchain technology in enhancing good governance transparency and development among African countries in line with Agenda 2030, *2018 Research and Innovation Conference*, 2018.

Pérez-Chacón, R. Asencio-Cortés, G. Martínez-Álvarez, F. and Troncoso, A., Big data time series forecasting based on pattern sequence similarity and its application to the electricity demand, *Information Sciences*, vol. 540, pp. 160-174, 2020.

Sellami, M. Mezni, H. and Hacid, M. S., On the use of big data frameworks for big service composition, *Journal of Network and Computer Applications*, vol. 166, 2020.

Shamim, S., Zeng, J., Khan, Z., and Ul Zia, N., Big data analytics capability and decision making performance in emerging market firms: The role of contractual and relational governance mechanisms, *Technological Forecasting and Social Change*, vol. 161, 2020.

Song, L., Zhang, K., Liang, T., Han, X., and Zhang, Y., Intelligent state of health estimation for lithium-ion battery pack based on big data analysis, *Journal of Energy Storage*, vol. 32, 2020.

Song, X., Wu, Y., Ma, Y., Cui, Y., and Gong, G., Military Simulation Big Data: Background, State of the Art, and Challenges, Mathematical Problems in Engineering, vol. 2015, pp. 1-20, 2015.

Suoniemi, S., Meyer-Waarden, L., Munzel, A., Zablah, A., and Straub, D., Big data and firm performance: The roles of market-directed capabilities and business strategy, *Information & Management*, vol. 57, no. 7, 2020.

Tang, J., Ma, T., and Luo, Q., Trends Prediction of Big Data: A Case Study based on fusion Data, *Procedia Computer Science*, vol. 174, pp. 181-190, 2020.

Valencia-Parra, Á., Varela-Vaca, Á., Parody, L., and Gómez-López, M., Unleashing Constraint Optimisation Problem solving in Big Date environments, *Journal of Computational Science*, vol. 45, 2020.

Viceconti, M., Hunter, P., and Hose, R., Big Data, Big Knowledge: Big Data for Personalized Healthcare, *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1209 – 1215, 2015.

Wang, S., Tsai, Y., and Ciou, Y., A hybrid big data analytical approach for analyzing customer patterns through an integrated supply chain network, *Journal of Industrial Information Integration*, vol. 20, 2020.

Wilkin, C., Ferreira, A., Rotaru, K., and Gaerlan, L., Big data prioritization in SCM decision-making: Its role and performance implications, *International Journal of Accounting Information Systems*, vol. 38, 2020.

Yang, J., Li, X., and Huang, S., Impacts on environmental quality and required environmental regulation adjustments: A perspective of directed technical change driven by big data, *Journal of Cleaner Production*, vol. 275, 2020.

Zhu, Y., and Jiang, Y., Optimization of face recognition algorithm based on deep learning multi feature fusion driven by big data, *Image and Vision Computing*, vol. 104, 2020.

Biographies

Mary Jane C. Samonte has a double bachelor's degree in computer education and information technology. She also has two post graduate degree; Information Technology and Computer Science. She finished her Doctor in IT with a study focusing in Deep Learning. She has a wide range of research interests that are centered around educational technologies, gamification, mobile and ubiquitous learning, digital game-based learning, artificial intelligence in education, e-health, assistive technology, natural language processing, green computing and data analytics-based studies.

Johndale Alfred A. Julian is a full-time Data Engineer at a company based in Madrid and a part-time teaching faculty at De La Salle-College of Saint Benilde. He graduated Bachelor of Science in Information Technology with Specialization in Game Design and Development. He is currently pursuing his master's degree in Information Technology at Mapua University under the supervision of the co-author of this paper. His research interests revolve around Data Science, Data Engineering, and Machine Learning.