Identifying the Substantial Big Data (BD) Aspects to Improve Product Engineering Design

Elmira Mohebi and Yvan Beauregard
Department of Mechanical Engineering
École de technologie supérieure (ÉTS) University
Montreal, Quebec, Canada
elmira.mohebi.1@ens.etsmtl.ca, yvan.beauregard@etsmtl.ca

Abstract

This paper suggests what aspects in terms of Big Data (BD) and New Product Development (NPD) have to be considered by companies in order to make a benefit specified in the engineering design phase. The research is based on reviewing related papers to big data and big data analytics. The different dimensions extracted from the papers include paying attention to data analyst role, big data analytics applications, the collaboration of data analysts and engineers, and performing a comprehensive NPD approach so as to improve new product development and product engineering design. The research shows that companies can leverage external and internal knowledge by considering a comprehensive approach to NPD. Finally, an application methodology has been explored to be a guideline for firms in order to facilitate their big data process.

Keywords
Big data, big data analytics, data analyst, new product development, engineering design

1. Introduction

1.1. Big data

In recent years Big Data (BD) has been attracting noticeable attention all over the world. Big data as a term emerged in 2010 and after that many definitions have been given. Big data presents information properties which have characterizations such as high volume, variety and velocity and it needs special technology and analytical methods in order to be transformed into value (De Mauro et al. 2016). Many definitions are similar in terms of volume, variety and velocity characterizations. Volume, variety and velocity are considered as large amounts of data, different kinds of data such as structured, unstructured or different sources, and high pace of flowing data respectively (Flankegard 2017). Zhai, Ong and Tsang (2014) referred to dimensionality as another characterization of data in their definition which regards different features and variables. BD with its characteristics has exceeded the power of a company to analyze and process it in an appropriate time (Bi and Cochran 2014). Big data can be regarded as interactive low cost and multi-media rich information obtained from enormous connections (Tan and Zhan 2014). Big data is ubiquitous and because of its unique challenges is being focused in different researches.

1.2. New product development and engineering design

There are various definitions for new product development. It is regarded as a process of introducing a new product or an idea to the market. These ideas can bring something that never have been seen before or are related to improving the existed products, adjusting product in order to meet the clients' requirements, or resolving a problem. Management of product innovation is a noticeable success factor in companies. Perpetuity and prosperity of modern companies is heavily dependent on developing and improving new products . Engineering design is aimed to produce a product to customers. Studies showed that most companies must consider New Product Development (NPD) for their profitability and growth (Zhan et al. 2017). From researchers point of view a proper product development process should empower companies with a more effective roadmap, and make them adaptable (Sheu and lee 2011). Many studies have showed that paying attention to the customers' requirements in engineering design process increases the likelihood of surpassing rivals in the market (Dixon et al. 2016).

1.3. Big data analytics

Analyzing big data and analytics techniques can help companies to leverage the large amount of data around them. The pioneers of Big Data Analytics (BDA) in a consulting firm have exceeded their rivals (Pearson and Wegener 2013). One of the conducted researches by IBM in cooperation with the Economist demonstrated that firms with applying big data and using

analytics are 36 percent more probable to be succeeding rather than their rivals. Big data analytics is applied in several areas. For instance, in a research it deals with hospitality to specified customer satisfaction and their experiences (Bi and Cochran 2014). Steed, et al. (2013) showed that conventional climate analytical methods are not sufficient in order to cope with data volume and complexity, and suggested a visual analysis tool for simulating earth system. Although, data analytics methods can use different tools of traditional statistics, multivariate analysis and machine learning, the crucial factor for data usage optimization is transforming a diagnostic approach to a prognostic one where developing predictive patterns is dependent on a platform where acquiring and analyzing data can be automated (Benedetti and et al. 2019). Data mining is one of the substantial examples of advanced intelligent business analytics and also is identified as one of the top technologies (Amani and Fadlalla 2017). Data mining is a process involved in recognizing valid, novel and understandable patterns in data (Pujari 2001). It can also be defined as the process of acquiring and exploiting knowledge from large sizes of data in order to improve decision (Han et al. 2011). Thus, data mining is concentrating on leveraging data properties of a firm to make financial and non-financial benefits (Amani and Fadlalla 2017).

1.4. Data analyst

Apart from state-of-the-art technologies and analytics techniques, the role of human in analyzing and interpreting the data is not ignorable. Data-driven designing is revealed as a supplement or substitution for the previous information-driven methods (Li et al. 2019). Data-driven modeling applies new machine learning techniques to construct patterns that catch physical manners (Solomatine and Ostfeld 2008). It can be helpful in the cases there are an enormous extents of data and it is not simple to create sufficient information-driven methods, because of weak conceptualization of the physical event or creating a mathematical model for that. As a consequence, there is a need to incorporate data scientists and experts into product development process (Porter and Heppelmann 2014).

1.5. Research Problem

There must be an effective way to help companies utilize big data in order to support their NPD from engineering design perspective. With the effective use of big data, companies can come up with new ideas or investigating about their products, clients and marketplaces and this is a great stride for NPD. We are observing an ever-increasing amount of data due to the fast technology development that make devices to be connected and exchanging data, which is recognized as Internet of Things (IOT) (Li et al. 2019). The big data analysis can affect the strategic and operational business process in companies and consequently may help them to gain competitive advantages (Tan and Zhan 2017). However, a main challenge for firms is how they can harness big data in design and engineering fields such as design analytics and supply chain management (Dixon et al. 2016).

To date, no research congregates different issues related to big data in the field of engineering design of new product. This paper seeks to find important aspects of big data for new product development through investigating literature and find out how companies could leverage big data and make benefit by applying it in engineering design. The next section is dedicated to reviewing the most related papers to this subject and synthesizing them. The key words used to find the papers were big data, big data analytics, and new product development. Then after reading their abstracts the main papers were those that cover the different perspectives toward big data. As a result of this study, an application for big data in engineering design of new product is identified in order to be a guideline for companies. In the end, the conclusion of the research is presented.

2. Theoretical Foundation

Big data is a comprehensive attitude to handle procedure by analyzing 3Vs (volume, variety, velocity) so as to build practicable eye sights for evaluating performance and authenticate competitive advantages (Tan and Zhan 2017). In recent years, it has been proved that burst of data is an obvious trend and, in the case it is applied appropriately, can lead to better product development for companies. There are several kinds of data, such as graphs, sensors, videos, texts and so on (Zhan et al. 2017). Furthermore, due to the variety changes from old structured data to unstructured data, the complexity of data has increased (Zikopoulos et al. 2012). Collecting, storing and analyzing data need some requirements. For example, for collecting data, Internet of Things (IoT) makes it easier to gain data from several data sources such as internet, industrial networks, and sensors and so on for the cloud (Bi and Cochran 2014). In the case of data operation, cloud computation suggests trustworthy duties through deployment of cloud data centers.

BDA is not only a technique or a technology; it is a complete combination of tactics, research, people and marketing (Bi and Cochran 2014). Big data analytics is a set of methodologies and tools such as regression examination, neural network analysis, cluster analysis, decision trees and etc. (Davis 2014). Big data analytics need a set of tools to recognize related data and turning them into beneficial knowledge. Companies require reassessing their methods of data management strategically and culturally in order to choose the proper data, and make correct decisions by that (Bi and Cochran 2014).

The effective application of big data can suggest many profits such as huge expenditure decrement, reducing the time for performing a computing task, increasing customer satisfaction, and new innovation and improvement in products and services. Nowadays, firms tend to perform simultaneous processing through cross-function groups and, therefore, this can lead to pace

improvement, efficient and flexible NPD process (Tan and Zhan 2017). Also, topical NPD approaches consider strong connection with customers in order to gain better understanding of customers' feedback for ongoing improvements. Companies are in touch with suppliers and competitors as well as customers in order to create comprehensive networks and reach competitive advantages.

In the following, the major investigated articles will be reviewed which help incorporating important aspects related to big data into an application.

2.1. The importance of data analyst role

Bag (2016) reported that big data and predictive analysis are critical factors for sustainability of future business. The need for having data scientists in this dynamic business is increasing so as to manage big data, acquire significant outcomes and improving management judgments. It is crucial for companies to choose and hire the appropriate data scientist in terms of their expertise set and span of analytics. Therefore, an effort has been made in order to identifying which skills are substantial for data analysts. The process includes selecting the analyst with skills in big data and analytics among three candidates based on eight skills and by three decision makers. The eight skills include technical knowledge, flexibility, time management, business intelligence, intellectual curiosity, strong interpersonal skills, supplier relationship management, and demand planning.

Finally, it concludes that three skills; intellectual curiosity, technical knowledge and business acumen are the most important skills for a data analyst. In the case of technical knowledge, basic statistics, knowledge of machine learning, scripting language (Python, Matlab), querying language (Sql, Hive, Pig), spreadsheet (Excel) and statistical language (R, SAS, SPSS) are important. Intellectual curiosity means having reasoning, rational attitude to problem, and the power of problem solving. Business acumen involves understanding the market for improving the company's lucrativeness.

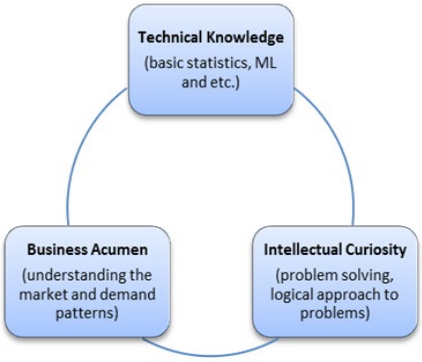

Fig. 1.Data analyst required skills in management

2.2.The collaboration of engineers and data analysts

Li et al. (2019) believe that today's smart products consist of intelligence as well as physical components. Data and the potential to process it into information and acts are substantial parts of products and the process to develop products. Likewise, a beneficial cooperation between engineers for the physical concept and data scientists for the data concept is significant in the NPD process in order to eliminate reworking and creating products with better physical and analytical components. At first, NPD process patterns for engineering and Knowledge Discovery and Data Mining (KDDM) approaches for processing data are investigated and then, incorporating KDDM into NPD is discovered in order to design the new model. Thus, the interplay models between engineers and data analysts must be recognized so as to improving deciding process and architecture model for the development of the new smart product.

The first step is identifying engineers' tasks and data scientists' tasks separately. The major engineering tasks in the NPD consist of inquiring possibility of product notions, improve manufacturing model notions, create empirical archetypes, evaluate production costs and discern manufacturing possibility (Ulrich and Eppinger 2012). For data analysts, these duties consist of gathering data, describing data, exploring data, and verifying data. There are also some tacit actions such as hypothesis designing and examining (descriptive analytics) and detecting data mining situations (predictive analytics). Overall, as you can see in *"Fig. 2"*, the engineering tasks are classified as identify and design concept, create and examine notions and assess them for election; likewise, the part of data science tasks are recognized as detect plus gather data, descriptive analyzing, confirm data qualification, and survey analyzing notions.

The Li et al., (2019) NPD3 model is shown in *"Fig. 3"*. The top section concentrates on the engineering activities for physical product design that is regarded as engineers' group responsibilities. They convert the client requirements into the technological specifications and then reach the optimal solution for the physical design. The lower lane shows the data knowledge group, who concentrate on processing data and analytic activities for developing the data product. They convert customer requirements into the data characteristics and use the data specification to reach the optimal data analysis solution. The middle lane shows the project management (PM) group that their responsibility is pursuing a stage-gate-based new product development procedure. The collaboration between engineers and data scientists is brokered by PM team. Engineers and data

scientist need disciplinary design information and it is crucial for straight interplays and messages exchange between these two teams. Thus, a project management team mediates to facilitate the collaboration across the team.

2.3. The importance of big data analytics tools and techniques

The Bi and Cohran (2014) surveyed about Big Data Analytics (BDA) and considered it as a crucial technique and technology to help data acquiring, storing and analyzing in data administration systems in up-to-date production. They studied about BDA applications to identify the requirements of production systems and also crucial enabling technologies for BDA. They specifically surveyed about BDA application in manufacturing systems. A production system requires knowledge systems to decide about system functions at various levels and areas.

The progression of a production system can be evaluated by size, intricacy and response ability of automation (Bi 2011). The evolvement of production technologies is categorized into the phases of applying Numerical Control (NC), Flexible Manufacturing Systems (FMSs), Computer Integrated Manufacturing (CIM), Distributed Manufacturing, and Predictive Manufacturing (PM). Common software instrumentations in order to help the production techniques include Quality Control (QC), Enterprise Requirements Planning (ERP-I), Total Quality Management, Product Lifecycle Management (PLM), Enterprise Resources Planning (ERP-H-II), Software as a Service (SaaS), and Platform as a Service (PaaS), Infrastructure as a Service (IaaS). The data dealt with knowledge systems has risen from flowing data in the digital age to BD in terms of volume, variety and velocity. There is a need for IT hardware systems to be compatible with converting data in an appropriate time. The computation circumstances have progressed from Microchip, mainframe, servers, and the Internet, to current cloud (Bi and Cochran 2014).

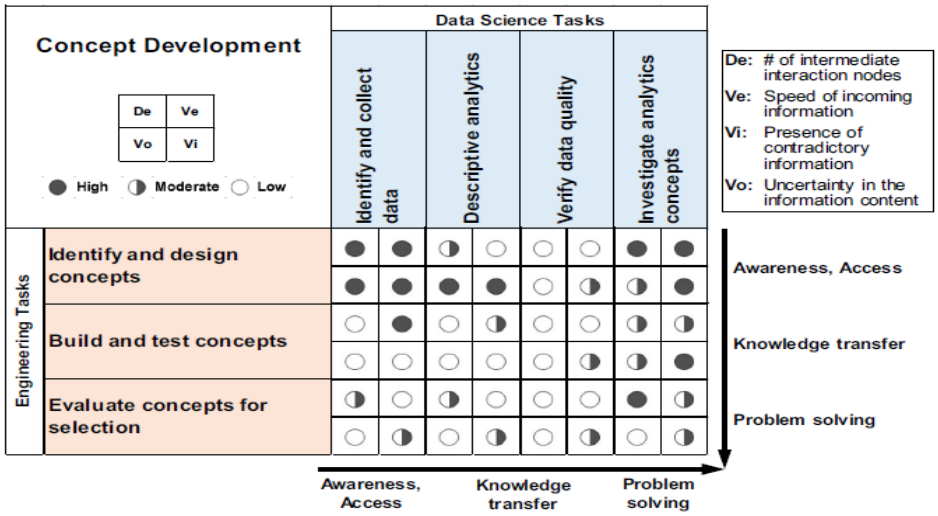

Fig. 2. Interaction models and attributes between engineers and data analysts. (Li et al. 2019)

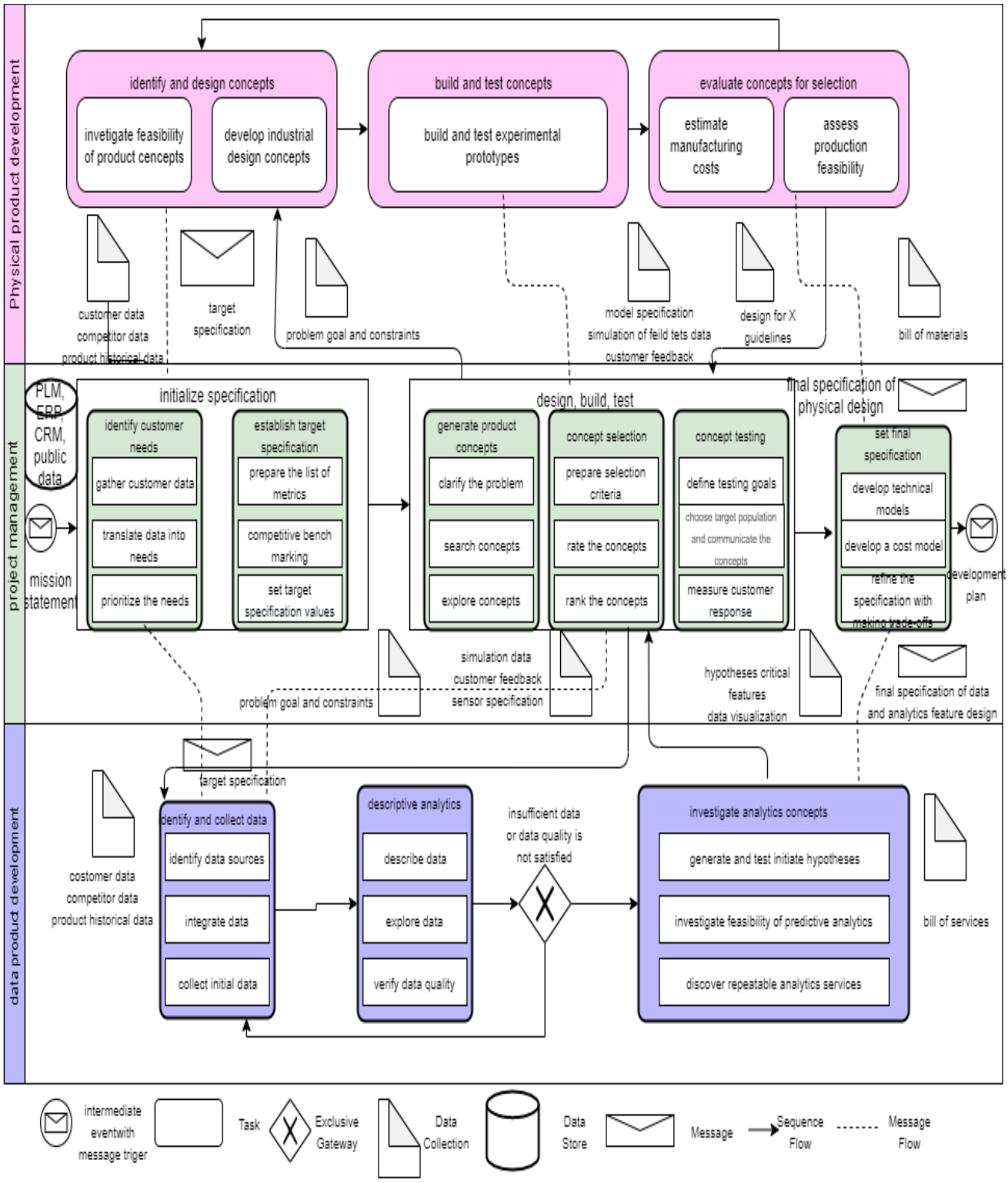

Fig. 3. An unified process model for NPD with data-diven attributes (NPD^3) (Li et al. 2019)

Improvement trends of enterprise system show that future systems need to be more collaborative, service oriented, responsive and predictive (Cooperative Synergies Inc. 2009). For example, in the field of new product development, managers through the use of the most important BDA tools such as text mining, data mining and sentiment mining try to improve new product strategies and they are aware of the social consumers' behaviors (Yu and Yung, 2016).

The BDA applications have some technical challenges. The crucial point for appliers is having a transparent recognition of the needs of big data applications, the ability of big data analytics, and the requirements of deployment. For example, there was a need to improve progressive platforms to supporting the requirements for real-time cloud based and light-weight practices. Today's firms require to gain a great deal of data from different datasets and also to apply the knowledge by Big Data Analytics (Doulkeridis and Norvag 2014). Modern analyzing platforms need to be involved with enormous measurable data, help low-latency data, and speed up progressive analysis designing and operating. Thus, BDA tools should consider the increasing amount of demands, size of data, type of users and computing load. The moving factor for big data is the software and platforms for foundation and analysis. Some preliminary techniques to speed up processing for large data sources include grid computing, in-database processing, in-memory analytics, and Hadoop (SAS, 2014). The first one prepares running workload adjustment and simultaneous operating for data analysis and data handling. The second one applies measurable design to decrease the processing time for data analytics. The third one deploy simultaneous, in-memory, multi-use disposals to data and quickly performs novel actions. Hadoop is a primitive infrastructure applied to dispense, catalog, handle and query data across multiple, horizontally scaled service nodes.

Overall, extracting knowledge from large amounts of data require scalable and smart algorithms, sets and applications. To this end, BD and its techniques are developed. The prosperity of a production organization is dependent on progression of IT to augment and elevate the value flow. Big Data Analytics support manufacturing firms to catch business opportunities, rapidly adjust to change and get along with uncertainty immediately. The common challenges for various BD applications are deploy, the tuning, and the further improvement of operating platforms accordance with a proper recognition of related use cases and work-loads (Hegeman et al. 2013). Furthermore, the achievement of a BD project is dependent on firm's culture and users' expertise.

2.4. Identifying the key success factors for facilitating product innovation process

Zhan et al., (2017) performed a BD framework to improve firms' product innovation procedure. After surveying previous studies the key success factors for product innovation were determined. These factors include: an accelerated innovation process, customer connection, and an ecosystem of innovation.

The accelerated innovation process tries to make innovation process faster through systematic methods. This factor helps companies gain competitive advantage due to the fact that they are the first on the marketplace which leads to faster recognition of profit. Also, the risk of changing the market before product launch is lower (Steinfeld and Beltoft 2014). Autonomy management considers the power of making decision on basic issues for employees; it can increase job satisfaction and motivation to work. Cross-functional teams decrease the manufacturing period time for NPD (Cooper and Kleinschmidt 2011). Simultaneous development is a way of designing and improving products that different stages perform at the same time. Using this method can lead to big benefits for firms, for example reducing program costs, potential risks, manpower requirements, and improving quality and flexibility (Lovelace et al. 2001).

Connecting with customers and understanding them is a substantial factor which can affect market share, revenue and margins positively, if their feedback is implemented in product. Customer connection considers four important elements; marketplace direction, customer relationship, understanding of customers, and having good communication with customers. Therefore, organizations pay more attention to customers and see them as sources of information and knowledge.

Ecosystem of innovation refers to creating an innovative surrounding that can help companies to develop new products in a high speed and low expenses way. This factor consists of five items; relationship with clients and associates, specialization of marketing examinations, quick expansion and setup, fast reaction to markets, and market and partner examinations. The ecosystem allows firms to produce greater value than a single company (Bogel et al. 2014). It is especially more beneficial when accompanied by information technology due to the cost reduction of coordination.

To sum it up, in *"Fig. 4"*, the most important items in accelerated process are: cross-functional teams work in parallel, autonomy management, divide projects into small elements, big data supported communication and development. The most important elements in connection are: understand customers and market via big data, keep close to customers, interact with customers and gather useful feedback, and update the latest product information quickly. The most important elements in ecosystem are: build strong networks with partners and customers, launch

the new product quickly, redevelop the product to meet the needs, gather feedback quickly from customers and partners.

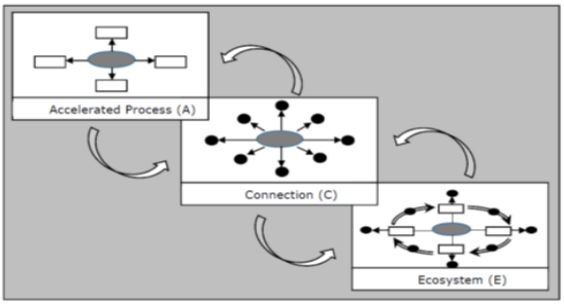

Fig.4: ACE Framework for product innovation
(Zhan et al. 2017)

3. Application Methodology

According to the reviewed studies, improving new product development in the field of engineering design through big data necessitates to consider BD from various aspects.

Firstly, it is important to provide the requirements in organizations for big data analytics. These requirements include considering data analysts skills and attributes compatible with the BDA work, preparing appropriate BDA infrastructures, tools and techniques, providing the proper situation for collaboration between persons responsible at BD. Secondly, companies have to pay attention to perform a comprehensive model, in which all the influential phases for improving NPD are included. After implementing and considering these steps, the outcomes from new product development will be visible. As shown in *"Fig. 5"*, the most important elements in order to leverage big data for NPD are gathered.

The following methodology has been explored as a guideline for firms to make them aware of their big data process especially for the engineering design phase.

3.1. Collaboration between engineers and data scientists

We should consider that creating new products needs to develop physical products and data products. In other words, engineers require cooperating tightly with data analysts to make decision about designing the product in order to support more data-driven attributes. Engineers as well as data scientists need cross-disciplinary design information and it is crucial to exchange direct interactions and messages between these two teams. Also, a project management team can mediate to help collaboration across the team. The time and context of exchanged information between groups improve knowledge sharpness and access, information transition and problem solving.

3.2. Regarding the role of data analysts

In many companies leveraging big data in order to improve new products is unattainable due to the lack of employees with the right skills and expertise. These skills may be different in different organizational contexts. According to previous studies, the most influential criteria and skills include technical knowledge, business acumen, and intellectual curiosity. Companies can hire individuals with the right expertise or can also train their employees through investing in learning programs.

3.3. Big Data Analysis

We should bear in mind that BD is not only about integration of data but also completely recognize the significance of data storage and processing data. Vigorous strives are required to improve progressive resolving from data attainment to embodiment, in order that the whole procedure can be implemented fluently within an appropriate time. Due to the big data tendency to increase in size dramatically, data analytics deal with big turbulences because of strong needs on computing, pace and diversity. The significant challenges of BDA include: improving storage systems to incorporate enormous sections of data from locative dispersed originals and performing manifold computing on the datasets. Measurable divided repository systems and performance machines are being improved to diminish these challenges.

Therefore, companies should take into consideration BDA applications capabilities and technologies compatible with the type and attributes of the data that they are dealing with.

3.4. ACE model

Considering a model and using big data in order to attain and join to a broad area of networks in every phase of product development lead to competitive advantage for firms. By performing the framework, novelties are built from inter-connected networks and these make companies able to immediately consolidate significant response from clients and partners. Project teams can revise the product process with regarding marketplace needs so that stay ahead of the competition. Companies can be more responsive to today's flowing variable information and evolving market circumstances with the innovation ecosystems.

This model is adopted from research conducted in 2017 with the aim of facilitating the product innovation process through a big data framework. The authors in this research have tried to extract all the key success factors regarding product innovation by doing a systematic literature review. In the next step, they interviewed the experts to improve and finalize identified key success factors, and at last, based on the previous steps, a big data framework has been evolved in order to improve product innovation and development in an organization. One of the advantages of this model is that all the success factors without any prejudice have been collected in the first step which includes four key success factors and 19 elements. Also, key success factors were evaluated and validated again through interviews by academics and experts in different industries. Another advantage of this framework is considering the use of big data in every phase of the product development that helps to accelerate innovation and also identifying product weaknesses in early phases. Besides, it has an impact on decreasing the cost and time of launching products and also identifying the key features and functionalities in customers' opinions. The major limitation of this framework is the lack of empirical studies and the fact that it needs to be tested over different industries (Zhan et al, 2017).

We continued reviewing papers in order to cover the framework limitations and assure about considering related factors through new related papers. An empirical research has been done in 2020 in order to examine how big data in different processes of new product development can affect the innovation performance. They considered five separate processes for a new product project that include; idea generation, business analysis, product design, test, and marketing. In order to test their hypothesis in which one of them was about the effect of big data on product design, they gathered project-level data from NPD projects in the USA, UK, and Australia in five different industries. Based on those data they proposed a model that inquires into how BD-embedded new product development processes. According to the findings, the effect of BD-embedded in the phase of product design is proved on the innovation performance in all three countries (Wang and Zhang, 2020). It is noteworthy that in a research with the aim of investigating the customers' agility role in NPD success, the authors found out that there is a connection between customer agility, data analysis tools and new product success. In which customer agility refers to how a firm understands and reacts rapidly to opportunities in the market and gain competitive advantage (Hajli and et al, 2020). Therefore, this can be mentioned as a validation for the importance of connection to customers. In terms of accelerated process in the ACE model, it also has been proved and mentioned as an outcome of using big data in the process of NPD in other papers. For instance, in 2019 a research in the food industry demonstrates that using BD in the developing of new product has an influence on reducing the time and cost without any effect on the products taste (Jagtap and Duong, 2019). Recent research that sheds light on BDA examines the combination of contextual factors in gaining process innovation. The positive point in this research is about including environmental and organizational factors and the authors believed that the combination of contextual factors can be varied for each organization. In this framework, the contextual factors are divided into three groups including BDA resources, environment, and organizational factors. BDA resources include data, technology, basic resources, technical skills, managerial skills, organizational learning, data-driven culture, and environment items are dynamism, heterogeneity, and hostility and the third group organizational factors include organization size and industry. Therefore, for each firm in accordance with the conditions and the aim of the innovation process, one or more combinations of factors can be influential than others (Mikalef and Krogstie, 2020).

4. Conclusions and Limitations

Data attributes play a major role in the success of a product as well as the physical specifications. This study refers to the crucial duty of big data in supporting companies to augment NPD specifically in engineering design of a product. There are many ways for companies to utilize big data and turn it into knowledge in order to expand their competence. We tried to investigate the literature in the areas of BD and NPD to deepen in BD and its different aspects in business. We found that firms should take BDA applications and their requirements, data analysts, and a comprehensive model for NPD into consideration so as to make a profit from big data. Utilizing big data in new product development necessitates firms to have employees aware of the data and with related data expertise, to

consider the proper way of analyzing data and the most appropriate application for that, to provide the proper environment for collaboration, and to consider the importance of NPD principles. Therefore, individuals have to pay close attention to different big data dimensions in their companies and understand all the theoretical concepts in order to apply them effectively in practice.

Aforementioned discussions can help companies improve their NPD. Firstly, it helps to specify the product weaknesses earlier in the development cycle. Secondly, the products can be accompanied with features and functionalities that are noticeable for customers. Thirdly, deleting unimportant features from the customers' points of view. Another point is identifying customers' needs priorities in markets.

The important point to mention is that the concentration of this paper is to inquire the important aspects of employing big data in engineering design. And also, the real firm situation is more intricate, because different organizations have various targets, research and development focus, BD technology, and accessible data and etc. As a suggestion, performing the methodology for a case study in the future will be fruitful to be more precise and assess the elements in practice.

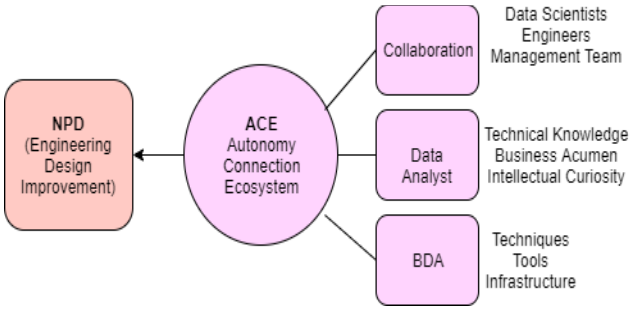

Fig.5 An approach for NPD by leveraging BD

References

Amani, F. A., & Fadlalla, A. M. (2017). Data mining applications in accounting: A review of the literature and organizing framework. *International Journal of Accounting Information Systems*, *24*, 32-58.

Bag, S. (2016). Fuzzy VIKOR approach for selection of big data analyst in procurement management. *Journal of Transport and Supply Chain Management*, *10*(1), 1-6.

Benedetti, A., Khoo, J., Sharma, S., Facco, P., Barolo, M., & Zomer, S. (2019). Data analytics on raw material properties to accelerate pharmaceutical drug development. *International Journal of Pharmaceutics*, *563*, 122-134.

Bi, Z., & Cochran, D. (2014). Big data analytics with applications. *Journal of Management Analytics, 1(4),* 249-265.

Bögel, S., Stieglitz, S., & Meske, C. (2014). A role model-based approach for modelling collaborative processes. *Business Process Management Journal, 20(4),* 598-614.

Bi, Z. M. (2011). Revisit system architecture for sustainable manufacturing. *Journal of Sustainability, 3(9*), 1323-1340.

Cooper, R. G., & Kleinschmidt, E. J. (2011). New products: The key factors in success. Marketing Classics Press.

Davis, C. K. (2014). Beyond data and analysis. Communications of the ACM, 57(6), 39-41.

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. Library Review, 65(3), 122-135.

Dixon, A., Liu, Y., & Setchi, R. (2016, August). Computer-Aided Ethnography in Engineering Design. In ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers Digital Collection.

Doulkeridis, C., & NØrvåg, K. (2014). A survey of large-scale analytical query processing in MapReduce. The VLDB Journal—*The International Journal on Very Large Data Bases, 23(3),* 355-380.

Flankegård, F. (2017). The use of data within Product Development of manufactured products.

Hajli, N., Tajvidi, M., Gbadamosi, A., & Nadeem, W. (2020). Understanding market agility for new product success with big data analytics. *Industrial Marketing Management*, 86, 135-143.

Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.

Hegeman, T., Ghit, B., Capotă, M., Hidders, J., Epema, D., & Iosup, A. (2013, October). The btworld use case for big data analytics: Description, mapreduce logical workflow, and empirical evaluation. In 2013 IEEE International Conference on Big Data (pp. 622-630). IEEE.

Jagtap, S., & Duong, L. N. K. (2019). Improving the new product development using big data: A case study of a food company. *British Food Journal*.

Li, Y., Roy, U., & Saltz, J. S. (2019). Towards an integrated process model for new product development with data-driven features (NPD 3). *Research in Engineering Design*, *30*(2), 271-289.

Lovelace, K., Shapiro, D. L., & Weingart, L. R. (2001). Maximizing cross-functional new product teams' innovativeness and constraint adherence: A conflict communications perspective. *Academy of management journal, 44(4)*, 779-793.

Mikalef, P., & Krogstie, J. (2020). Examining the interplay between big data analytics and contextual factors in driving process innovation capabilities. *European Journal of Information Systems, 1-28.*

Pearson, T., & Wegener, R. (2013). Big data: the organizational challenge. Bain Co.

Piengang, F. C. N., Beauregard, Y., & Kenné, J. P. (2019). An APS software selection methodology integrating experts and decisions-maker's opinions on selection criteria: A case study. *Cogent Engineering*, *6*(1), 1594509.

Porter, M. E., & Heppelmann, J. E. (2014). How smart, connected products are transforming competition. *Harvard business review, 92(11)*, 64-88.

Pujari, A. K. (2001). Data mining techniques. Universities press.

SAS (2014). Big data meets big data analytics: Three key technologies for extracting real-time business value from the big data that threatens to overwhelm traditional computing architectures. http://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/big-data-meetsbig-data-analytics-105777.pdf.

Sheu, D. D., & Lee, H. K. (2011). A proposed process for systematic innovation. *International Journal of Production Research*, 49(3), 847-868.

Solomatine, D. P., & Ostfeld, A. (2008). Data-driven modelling: some past experiences and new approaches. *Journal of hydroinformatics, 10(1)*, 3-22.

Steinfeld, E. S., & Beltoft, T. (2014). Innovation lessons from China. *MIT Sloan Management Review, 55(4),* 49.

Tan, K. H., & Zhan, Y. (2017). Improving new product development using big data: a case study of an electronics company. *R&D Management*, *47*(4), 570-582.

Troester, M. (2012). Big data meets big data analytics. Cary, NC: SAS Institute Inc, 1, 5.

Ulrich, K. T., & Eppinger, S. D. (2012). Concept selection. Product Design and Development, 5th ed. Philadelphia: McGraw-Hill/Irwin, 1, 145-161.

Vaughan, G. (2018). Efficient big data model selection with applications to fraud detection. *International Journal of Forecasting*.

Wang, Y., & Zhang, H. (2020). Achieving Sustainable New Product Development by Implementing Big Data-Embedded New Product Development Process. Sustainability, 12(11), 4681.

White, T. (2012). Hadoop: The Definitive Guide, 3. Aufl., Beijing. Zikopoulos, S./Eaton, C.(2011): Understanding big data: Analytics for enterprise class hadoop and streaming data.

Yu, Shidong, and Dongsheng Yang. "The role of big data analysis in new product development." *2016 International Conference on Network and Information Systems for Computers* (ICNISC). IEEE, 2016.

Zhan, Y., Tan, K. H., Li, Y., & Tse, Y. K. (2018). Unlocking the power of big data in new product development. *Annals of Operations Research*, *270*(1-2), 577-595.

Zhan, Y., Tan, K. H., Ji, G., Chung, L., & Tseng, M. (2017). A big data framework for facilitating product innovation processes. *Business Process Management Journal, 23(3),* 518-536.

Zikopoulos, P. C., & Eaton, C. (2012). deRoos D., Deutsch T., Lapis G. Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data, 1.

Biography / Biographies

Yvan Beauregard is a Professor, in the Department of Mechanical Engineering at the École de technologie supérieure (ÉTS) University, Quebec, Canada. He earned B.S. from Polytechnique Montréal University, Masters in MBA from Mcgill University, and Ph.D. from Concordia University of Montreal. He has published journal and conference papers. His research interests include aerospace, information and communications technologies, and health and his research areas include project management, operation management, quality management, product development and lean engineering.

Elmira Mohebi. is a student of Project Management Engineering at École de technologie supérieure (ÉTS) University and holds a Bachelor of Economics from Razi University and a Master of Business from SBU University in Iran.