CRITICAL ANALYSIS OF PROJECT MANAGEMENT DURING THE USAGE OF BIG DATA

UNIVERSITÁ DEGLI STUDI DI MILANO

FACOLTÁ DI SCIENZE POLITICHE, ECONOMICHE E SOCIALI

Laurea Magistrale in Data Science and Economics

Author: Pramodh Chinthareddy,
Matriculation Number: 936461

Supervisor: Prof. Maria Laura Toraldo,
Dept. Of Economics,
Management and Quantitative Methods

Co-Supervisor: Prof. Silvia Salini,
Dept. Of Economics,
Management and Quantitative Methods

Academic Year 2019-2020

I declare that this material, which will be now submitted for assessment, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work. I understand that plagiarism, collusion, and copying are grave and serious offences in the university and accept the penalties that would be imposed should I engage in plagiarism, collusion or copying. This report, or any part of it, has not been previously submitted by me or any other person for assessment on this or any other course of study.

Abstract

We live in an electronic, digitally mediated age. New technologies and new sources of data have emerged in the past few years, particularly in relation to electronic data. Analysis of everyday resources which we use in making our observations can have a huge impact on resource allocation and overall productivity.

Big data helps in increasing speed of project intelligence that can enable effective analysis through which project managers can use business intelligence tools to execute projects. Through the aid of big data, project managers can maintain their working principle regarding projects that can be helpful for project members.

First chapter comes up with aim to find the impact of big data analysis in project operations to maintain accuracy and flexibility. It can help in identifying actual requirements for handling big data applications and analysis to improve efficiency in resource and project management. It can also develop knowledge about criteria of big data analysis within project activities.

Second chapter is a literature review of the impact of Big Data and other relevant technologies towards Project Management across varied industries and fields. It's an extensive study in the fields of Manufacturing, Supply Chain, where automated models were utilized for process handling and decision making.

Third chapter aims to analyse impact of suitable selection of procedures for information collection regarding big data in improvising management in companies. It identifies Qualitative research methods followed by instrumentation that aid in meeting research aims and objectives. 3 Project Managers and 1 Machine Learning Engineer from different companies were interviewed to better understand about projects, their execution and the Big Data models as revenue generating platforms.

Fourth chapter concludes the study by interpreting information from the Qualitative data gathered and how the study addressed its Research Objectives and Research Questions.

Acknowledgement

I would like to thank my Thesis supervisor Prof. Maria Laura Toraldo and co-supervisor Prof. Silvia Salini for their guidance and support in finishing this research study.

I would like to thank the Project managers, Product managers and Machine Learning Engineers who took part in the Interviews and shared valuable insights.

Finally, I am thankful to my parents, friends who helped me in every sphere of life. Their contribution really helped me to complete this study effectively.

Contents

List of Figures

Chapter 1

Introduction

1.1 Research background

The significance of big data helps in creating opportunities for growth and also provides a way of analysing industry data to maintain progression. Through big data, most of the companies can restore their existing databases by expanding storage facilities. It can help them to use those databases that can be beneficial for storage management systems. In project management, big data plays an effective role in providing a reliable structure due to which project managers can maintain project management processes. According to the report, in 2012, 58% of companies have agreed to use big data technologies for project management and in 2014, 73% of organisations have decided to use big data within their project assignment (FrankovÃ¡ et al., 2016). It can allow them to use agile methods to maintain project related activities that can help in providing accurate outcomes depending on situations.

Big data helps in increasing speed of project intelligence that can enable effective analysis through which project managers can use business intelligence tools to execute projects. Through the aid of big data, project managers can maintain their working principle regarding projects that can be helpful for project members. It can allow them to perform their project operations in a structured manner that can assist in getting positive outcomes through their project management operations. According to Zerbino et al. (2018), through big data analysis, general managers of projects can make proper decisions to organise each segment or phase of projects. It can assist in arranging all non-functional operations due to which stability and accuracy can be maintained successfully. It can help in storing a huge amount of data regarding projects within project management databases. Therefore, project members and managers can handle their decision-making process through big data analysis within project assignment.

General Managers are responsible for providing a proper structure of framework for project members that they can understand their roles to execute their project operations. Using big data, general managers can collect more and more information about marketing movement, current trends and several project related activities. Therefore, general managers can take adequate steps in arranging these requirements to accomplish their assignment within a fixed time frame. Big data also helps in providing accurate information about costs of marketing products or services towards managers. Therefore, they can reduce project costs depending on situations and they can predict future and current trends due to which reliability and accuracy can be developed within project management operations. Moreover, planning and other resource related events can be controlled with the help of big data analysis through which efficiency in performance can be achieved during operating tasks of projects. It can also allow project members to follow a significant way or structure to accomplish all kinds of tasks regarding project assignment.

Project efficiency is the key of arranging sustainability within a project assignment through which possible outcomes can be organised successfully. Through the assistance of big data, general managers are able to gather specific information about processes that can help in achieving goals of projects. Through big data analysis, general managers can easily identify risk factors that can affect directly towards project assignment. Therefore, general managers can make decisions through decision making processes to deal with any kind of situations that affect project growth. Moreover, optimal performance can be achieved through the assistance of big data due to which general managers are able to maintain the flow of project operations. Thus, project members are able to provide a structural way of managing project processes through which accuracy and efficiency in project activities can be developed by using big data. It can allow project managers to operate their work in an optimal manner to achieve objectives and goals.

Through the assistance of big data, general managers can gather specific information about resource management. It would help general managers to understand resource requirements to accomplish their tasks due to which budget and schedule related operations can be organised successfully. It can bring a stable environment for project managers to maintain possible outcomes with reliable resources using big data analysis. This analysis can also lead to bring appropriate solutions for resource management due to which general managers can change their existing services to improve sustainability within project assignment.

1.2 Research problem

Research problem represents rationale of research topic through which issues in performing project management operations can be identified. Big data is one of the key aspects for project management to increase efficiency and transmission speed. It can allow general managers to accomplish their project activities to reach their project objectives or goals. However, sometimes, due to complexities in big data program, general managers have failed to manage big data analysis during performing project activities. It can decrease efficiency and also slow the progression rate due to which effectiveness and other clarity in execution process cannot be organised properly. Due to issues, general managers can also fail to make proper decisions that can directly impact on outcomes of projects. Thus, general managers are expected to understand this situation and modify their existing services to manage big data analysis. It can help them to accomplish their project goals without any issues in big data analysis in this current world.

According to the report, in 2015, 60% of big data project management has failed to reach their destinations due to lack of proper knowledge in this field. Moreover, this report has also stated that about 83% of organisations have failed to use big data in 2017 in their project management due to lack of work strengths and abilities (Max Henrion, 2019). Due to these issues, general managers and project members have failed to execute their individual operations due to lack of understanding and failed to meet their targets within the given timeline. Due to lack of experiences in marketing environment, project managers and members have failed to analyse requirements of their project operations. It can decrease the efficiency and clarity in service operations within project management activities. Due to lack of knowledge in big data or Hadoop application, project members cannot utilise map-reduce program within their project operations. It can directly affect the outcome of projects due to which project members cannot organise any task related information within project operations. It can decrease the processing structure and also reduce the way of managing schedule and cost management processes.

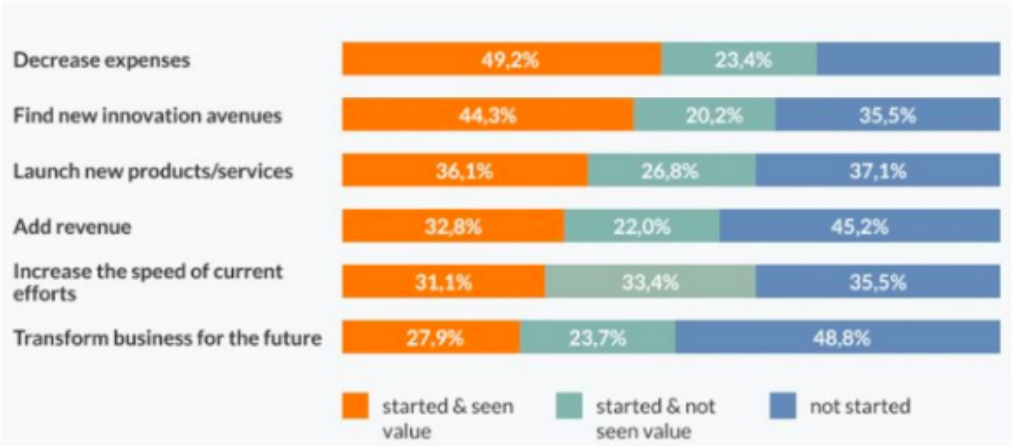

Figure 1.1: Ration of big data usage in project management (Source: As influenced by Larson, 2018)

On the other hand, due to lack of proper skills in Hadoop application or big data analysis program, general managers can fail to make proper decisions. According to Larson (2018), it cannot bring a stable environment due to which sequence within project operations cannot be executed properly due to these issues. Sometimes, using Hadoop, project managers and members have failed to map using Map and reduce program their project operations properly. It can directly affect planning process for project management operations and also slow the development process as well. These issues can also decrease the level of accuracy through which project managers cannot provide a reliable path for their members due to which they have failed to understand the way of managing projects. Due to lack of analytical skills, members of projects can fail to comprehend situations due to which general managers cannot use appropriate structure for their project operations. It can decrease the level of sustainability within project operations that can also slow the progression of project activities.

General Managers of projects are expected to take some responsibilities in analysing these issues and take proper measures to handle it. General Managers are needed to deliver accurate information about the usage of big data tools or Hadoop applications towards their members. Therefore, they can understand the usage of map-reduce program and utilise this program within assignment to increase transmission speed and also organise structure of project operations. As proposed by Wang et al. (2016), it can help in bringing a sustainable structure due to which project operations can be maintained in a synchronised manner. It can also allow members to follow a proper guideline to manage the sequence of every phase of projects through which complexities within projects can be avoided. It can increase clarity and stability regarding big data analy-

sis within project assignment.

1.3 Research aim

The aim of this research is to find the impact of big data analysis in project operations to maintain accuracy and flexibility. It can help in identifying actual requirements for handling big data applications and analysis to improve efficiency in resource and project management. It can also develop knowledge about criteria of big data analysis within project activities.

1.4 Research objective

- To determine the way of arranging project activities using big data analysis within a project assignment.
- To recognise issues of big data that affects the overall structure of an assignment within a project assignment.
- To recommend adequate strategies to manage issues or challenges of big data within a project assignment.

1.5 Research question

- What are the ways of big data in terms of managing project operations within a project management?
- What are the issues of big data that negatively impact processes of projects within a project assignment?
- What kind of strategies is helpful for resolving issues of big data during managing project phases?

1.6 Significance of research

Project management helps in arranging project tasks through which project managers can take proper steps to handle all activities with the help of big data tools. It can help in providing a way of managing resource, schedule and cost management processes through which efficiency and structure of projects can be organised properly. It can help project members to organise their requirements due to which they can deliver quality

services towards their assignment to deal with any kind of complexities. Through the assistance of big data, members of a project can analyse external situations and deliver accurate information towards their general manager. It can allow general managers to maintain flow of information due to which accuracy level can be developed without any issues. It can help in arranging possible outcomes due to which stability and structure within a project can be organised properly. It can assist in improving project efficiency and deal with any kind of circumstances without having any complex situations.

On the other hand, there are several issues that can directly affect the process of execution within a project assignment. It includes lack of knowledge, lack of analytical skills and lack of understanding. Due to these issues, project operators have failed to understand actual requirements and other essentials for their project that can slow the progression and also decrease clarity in service operations. Due to lack of knowledge, members have failed to organise their project operations due to which sustainability in project structure cannot be organised properly. It can also slow the process of development within project management that can also decrease accuracy within project assignments. Moreover, due to lack of understanding, members have failed to know the working principle of Hadoop application. Due to this issue, general managers have failed to guide their team members that can also negatively effect on progression and other issues within project assignments.

In project management, using big data analysis and tools, project managers can develop a way of gathering required information that can be beneficial for project assignment. Through this information, service operators are able to handle their individual tasks within project assignment. It can help in bringing a proper structure for project operations due to which stability and accuracy in managing every task within projects can be developed. Therefore, general managers can guide their members adequately so that they can understand their roles and they can give their potential performance to reach destinations. It can also help in avoiding internal and external issues within projects using big data analysis. Therefore, project members and general managers can develop accurate structure through which flexibility and sustainability within project operations can be organised.

Chapter 2

Literature Review

2.1 Conceptual Framework

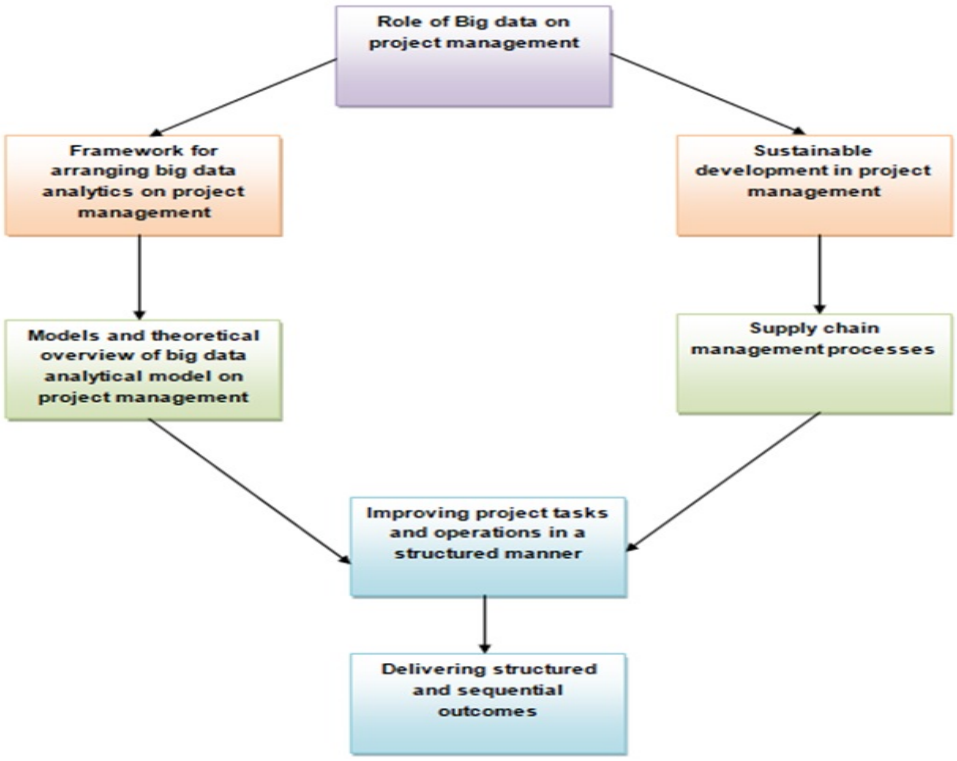

Figure 2.1: (Source: Created by author)

2.2 Manufacturing

This study helps in presenting theoretical knowledge of research study through which critical knowledge about content can be analysed properly. The purpose of this literature review is to identify crucial factors of big data analysis on project management while operating each task in a synchronised manner. This section explains about the importance of big data to maintain sequence and structure of project management through which effective operations can be maintained properly. It can also develop knowledge about the structure of knowledge production regarding big data and project management practices with the help of big data analysis. Through this process, processes of every task and theoretical overview of project management operations while using big data concepts in this modern world situation.

The significance of a cleaner production approach helps in improving lifecycle of project management coordination and optimisation process can be organised successfully. Through the usage of service driven and big data analytics patterns can guide project members to overcome any complex situations. It can allow in implementing a proper plan for project assignment through which project managers can easily gather knowledge and data about current situations. As influenced by Zhang et al. (2017), through the aid of CP strategy, project managers and team members can focus on project progression due to which stability and availability of resources can be managed properly. It can assist in improving operational tasks within project assignment and also deliver accurate results depending on requirements as well. Moreover, data mining plays an effective role in organising gathered information through which schedule and cost management processes can be structured in a synchronised way. Thus, project managers can make proper decisions with the assistance of their team members and senior management to reach objectives and goals.

Communication and sensor technologies provide a structured platform for project managers and members to organise advanced processing and simulation models. It can assist members to understand their roles and responsibilities to execute their tasks in a synchronised way. Thus, all selection and virtualisation operations regarding project management can be executed in a structured way. It can assist in developing the process and other manufacturing operations without having any complexities. However, sometimes, project members cannot utilise project activities and manufacturing operations due to complex environment and structure. As opined by Babiceanu and Seker (2016), it can decrease performance quality and accuracy in service operations within a project management while using cyber physical systems. Adoption of

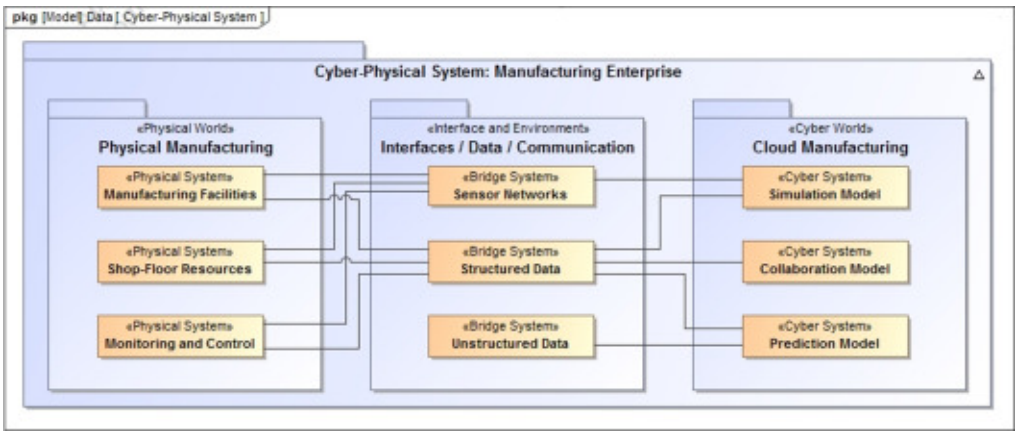

Figure 2.2: Cyber Physical Systems. (Source:https://www.sciencedirect.com/science/article/abs/pii/S0166361516300471)

the Internet of Things can improve manufacturing domain through which equipment of project operations can be modified according to requirements. Therefore, project members can gather specific knowledge and requirements through which stability and accuracy can be organised successfully. It can assist project managers to implement a proper plan for their execution process that enables effective solutions for project members. Therefore, they can easily execute their operational tasks in a structured way in transformation technologies. Supply chain management is one of the crucial aspects of project management operations through which service and manufacturing sectors can build a sequence of their project management operations. Through the usage of big data, project managers and members can gather specific data and information to maintain database management system in SCM. It can help service operators to understand their requirements for achieving success. According to Zhong et al. (2016), it can bring a proper structure for their service operations due to which accuracy and decision-making processes can be managed properly. However, due to some complexities in decision making process, project managers have failed to deliver quality information towards their project members. It can directly affect actual and planned progression due to which stability and accuracy in database management and decision-making process cannot be managed adequately. Apart from these, project managers are expected to take proper steps in understanding their issues and develop their decision-making process to improve performance structure. It can assist project members to use effective resources and improve their project management operations in a synchronised manner. Thus, they can easily achieve their goals and objectives without having complex situations.

2.3 Supply Chain

Big data allows project managers to understand the idea of organising production and innovation processes for their project assignment. It can give a secured and sustainable way of arranging value creation through which managing and interpreting capabilities can be managed properly. However, sometimes, due to lack of data analytical techniques, project managers would fail to analyse current situations that can directly impact the planned and actual progression of a project. Therefore, project managers are expected to modify the infrastructure of data analytics and organise each segment of project operations to organise data management processes. According to Tan et al. (2015), it can assist project managers to make appropriate decisions to maintain the flow of execution and accomplish all requirements. It can improve data set and generate all useful information about innovation in resource management processes using supply chain. Therefore, project operators can use those resources to develop their skills of their project operations.

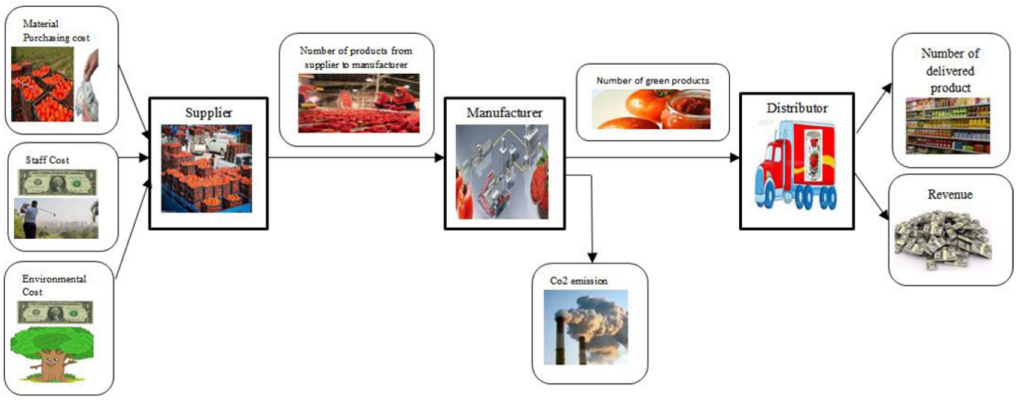

Figure 2.3: Structure of supply chain. (Source:https://www.sciencedirect.com/science/article/abs/pii/S0305054817301405)

Big data plays an effective role in managing performance evaluation through which supply change management operations can be executed in a structured way. Through Network DEA, efficiency and multi stage processes can be developed through which sustainability in project operations can be maintained properly. As followed by Badiezadeh et al. (2018), a big data approach can assist in increasing performance clarity through which verification and value of each segment in project management operations can be maintained properly. On the other hand, due to some complexities, some project managers and project members cannot acknowledge the way of managing project operations. It can directly affect the project progression due to which stability and accuracy cannot

be executed effectively. Through the assistance of DEA and NDEA models, issues regarding supply chains and project tasks can be maintained properly. It can allow project managers to understand every task through which effective measures can be organised properly. It can develop the process of executing all operations in avoiding complexities and bring a sustainable atmosphere. Therefore, project operators can give their effective performance to achieve their goals without having any complexities. Descriptive analytics and network analytics are two crucial aspects for managing big data analysis through which supply chain practice and other operations can be maintained properly. It can assist in improving performance and delivery processes through which risk factors and other operational tasks can be maintained adequately. Therefore, project managers and project operators can analyse the standard of environmental activities through which product development and other operations can be maintained properly. However, performance in project activities cannot be managed due to lack of resources that can negatively impact on project sequences. The project manager can easily analyse project operations and other activities properly to identify risk factors and take proper precautions to avoid those issues. Thus, project members can easily avoid complexities due to which structure in execution processes can be maintained properly. As followed by Chae (2015), using content and descriptive analysis, structure of database management operations can be maintained successfully through which distribution processes can be maintained properly. It can improve the way of managing significance of project operations and also deliver quality outcomes depending on requirements. Decision making abilities play an effective role in arranging all uncertainties and risk factors through which real attributes can be organised properly. Through transformation and developing processes of decision making and other activities of project operations can be organised properly. It can give opportunities for project members to maintain the sequence of performing project operations by avoiding risks and uncertainties. Therefore, project members can follow big data analytics to control and diagnosis related operations can be maintained properly. It can allow project members to deliver quality requirements depending on customers' needs. According to Wu et al. (2017), sustainability indicators of big data can provide a way of arranging all motivation progression through which sustainable goals can be achieved adequately. It can bring a stable atmosphere due to which different sets of indicators can be maintained properly that can decrease the way of managing sequence of evaluation process within project assignment. Through the aid of decision-making process, project managers can easily fulfil all requirements through which sustainability and other operations can be analysed properly. It can allow in understanding potential opportunities and risks

due to which risk measurements and other project activities can be organised properly. It can also develop the way of managing clarity and efficiency within project assignment.

2.4 Project Management

Big data plays an effective role in shaping the future for project management operations through which project schedule and project operations can be organised in a sequential way. Through big data, project managers and project operators can use effective methods to store equivalent results and organise sequences of every task without having any complexities. Zerbino et al. (2018) asserted that development of project management can be feasible if members tend to regulate, monitor and control properly. As influenced by Larson, (2018), through the usage of big data, project operators can store appropriate information and improve production and project management practices through which issues and other complexities can be avoided adequately. As followed by Zerbino et al. (2018), using these programs, project members can easily store quality and both structured data (using relational databases such as SQL) and unstructured data (using NoSQL databases such as MongoDB) within the database management system to arrange the balance between project initiation and project planning phase. The project manager is expected to gather specific information about requirements of handling big data through which they can use those requirements for increasing speed of project management operations. It can decrease complexities and improve structure depending on requirements within project management. The project managers can easily implement their project plan according to their requirements for making benefits and projects can reach their destination while using big data. Special volume plays a crucial role in arranging nature resources for managing sustainable development for project operations. Big data helps in providing a way of analysing dynamic nature of sustainability and development for project assignment. Therefore, project managers can guide their team members to follow a sequence of managing production processes through which performance of projects can be developed. On the other hand, using an evidential reasoning approach, project managers can focus on every segment of their project operations through which goals and objectives can be achieved in a sequential manner. As opined by Song et al. (2017), using big data, natural resources of every task can be managed due to which resource management processes can be organised properly. It can assist in delivering quality essentials towards their members that they can easily achieve their targets in project as-

signment. Supply chain management processes can also be organised within the project assignment through which structure of every task can be organised successfully without having any complex situations. Data envelopment analysis and Decision-making process help in improving structure of relative performance for project execution operations. It can guide project managers to measure alternative methods that are beneficial for managerial efficiency through which achievement criteria can be organised successfully. On the other hand, project managers can analyse required improvement through these processes through which they can reconstruct a plan for arranging new schedules for their project assignment. As proposed by Iyer and Banerjee (2016), it can assist in improving structure due to which stability and accuracy in service operations within project management can be maintained properly. Therefore, project operators can follow a sequence of arranging sustainable operations without having any complex situations through these processes. Data collection and efficiency measurement approaches can also assist project managers to identify key requirements for arranging project management operations. It can bring stability in operations through which project operators can organise each and every segment properly without having complex situations. Communication is one of the crucial aspects in project management through which relationships between trust, clarity and communication can be organised properly. Project team members are expected to know their strengths and give their effective performance to achieve their goals. However, sometimes, team members have failed to understand each other due to lack of communication and trust among each other. As described by Henderson et al. (2016), it can negatively impact project progression and fail to deliver quality outcomes depending on requirements. Therefore, project managers and sponsors are expected to take proper steps in arranging a sequence of establishing a relationship between project managers and team members. It can assist in improving performance rate due to which clarity and sustainability can be organised properly without having complex situations in project management operations. Effective communication is required for sharing knowledge and adequate information through which sustainability and accuracy in service operations can be improved. Thus, project members can build trust and clarity among each other without having any complex situations in project management operations.The significance of project space model helps in developing understanding skills and communication process for project operators and project managers. Through the usage of project communication tools, project managers can easily provide accurate information towards their project operators that they can execute their tasks in a structured manner. It can improve the way of managing project aspects through which project managers

can easily understand the status of their project and use appropriate strategies depending on requirements. In the words of van der Hoorn (2016), it can provide confidence towards project members to develop their knowledge and skills to maintain the flow of project operations. However, due to lack of effective resources, project members cannot give their effective performance that can slow the progression and also the improvement process for project management. Apart from these, project managers can identify requirements of their project members using project space model to maintain the project life cycle through which stability and accuracy can be organised effectively. Thus, project operators can analyse required resources for their project operations and deliver that information towards their project managers. Therefore, project managers can take proper steps in fulfilling those requirements and manage project life cycle. Project risks help in explaining shortcomings of a project due to which growth and sustainability of a project cannot be maintained properly. Project managers are responsible for analysing project risks that can decrease the way of improving project management operations. Therefore, it can directly affect the sequence of project management operations and also slow the process of execution. Due to project risks, project members cannot acknowledge proper requirements for their project management operations. As proposed by Liu et al. (2016), it can decrease the process of managing all kinds of tasks within a project management. Therefore, project managers can take proper steps in analysing requirements to deal with any kind of tasks. It can bring a stable atmosphere for project management operations due to which improvement and quality can be maintained properly. Due to bugs in project management system, confidentiality and efficiency in project operations cannot be organised that can slow the progression progress by disclosing all information. It can decrease the quality of project processes that can negatively impact on security management processes within project management. Moreover, due to unexpected behaviour of database management system, all information regarding project operations cannot be organised properly. Thus, confidentiality and effectiveness in project management cannot be managed due to which structure of database management system within project cannot be organised properly. Therefore, project manager is expected to take proper actions such as improving security and managing network using P2P communication to manage these issues. Through the usage of risk dimensions, project members and project managers can easily analyse project performance and risk factors. Therefore, they can make accurate decisions depending on requirements to deal with any kind of situation to manage efficiency within project assignment. It can also develop the process of executing any kind of tasks and maintain the flow of project related oper-

ations within a project management. Change management is one of the crucial aspects of project management through which big data analysis procedures can be executed properly. The project manager is expected to take responsibility in reconstructing project management operations to increase knowledge about production.It can help project members to manage all sequences of project management operations due to which stability and accuracy in service operations can be maintained properly. As influenced by Bresnen (2016), it can allow in improving theoretical approach through which project members can easily develop their professional practices. However, implications in value added services in project management can decrease the process of project management practices. It can decrease credibility and accuracy in service operations. The project managers and members can take proper steps in resolving those implications and improve quality and structure through which stability can be managed properly. According to Bresnen (2016), through innovative thinking, project members can easily develop their sequence of project operations due to which possible outcomes can be gathered effectively. It can bring stability and develop accuracy in performance management due to which structure in execution processes can be maintained properly. Efficiency in project management can assist in improving structure of operational tasks within a project assignment. The project manager can analyse operational risks due to which project performance can be maintained properly. It can allow in analysing all activities and also maintain the flow of operational tasks. Thus, possible outcomes depending on requirements can be managed properly that can be beneficial for project management activities. Due to complexities in risk management processes, operators have failed to analyse sequences of their project operations that can directly affect project progression. As proposed by Paquin et al. (2016), using PEM and PRM, project managers can guide their members to understand each segment with proper information. Therefore, they can use effective resources to solve any kind of risk factors through which potentiality and accuracy level can be managed properly. Through the usage of these approaches or operations, project managers and members can easily manage threshold value and implement a proper plan for their project operations. It can bring a stable and complex free environment through which accuracy in project management operations can be executed properly. Therefore, performance and quality can be developed without any issues and assist in reaching goals for project assignment. Global software development process helps in analysing challenges and risk factors for project management through which successful criteria for a project assignment can be analysed properly. Using positive aspects of software development, project members can identify requirements of their service operations and also implement

a proper plan for arranging all tasks within their service management. As influenced by Niazi et al. (2016), it can assist in managing implementing processes through which success factors and other operations in project management can help in managing sequence of practical knowledge in project assignment. Project members can achieve their goals and objectives after maintaining processing structure of their project. It can guide project members to perform their operational tasks due to which accuracy and structure can be managed properly. On the other hand, project managers can use GSD to maintain sustainability and accuracy in service operations in service management. Thus, project managers and operators can easily reach their destination in this project management. It can assist in managing strategic implementation processes to reach objectives and goals using strategic movements. Project manager guides project members to follow a structured sequence according to the project plan to achieve goals. Using several strategies and plans, project operators can identify required essentials for managing each task through which efficiency and structure can be organised properly. Strategies include professionalism, external knowledge skills and technical abilities. Through these strategies, project manager can identify required essentials for project assignment. It can allow in managing all demands and flow of project operations. It can help in organising quality and developing structure of projects while using big data. Through project management plan, project members can give their effective performance to maintain the flow of operations through which operational tasks can be executed properly. As proposed by Hodgson and Paton (2016), it can increase the way of managing sequential outcomes that can provide required results according to the needs of current situations. It can assist in managing professionalism to improve standards within project management operations. Project managers are also responsible for making proper decisions depending on requirements of their project members. Thus, relationships between project managers and project members can be organised effectively without any issues. On the other hand, due to lack of fundamental knowledge, project members cannot identify required services due to which project managers cannot organise their project operations. Thus, they have failed to execute their operational tasks due to which they can make proper steps to deal with any kind of complexities. It can increase the way of managing professional status due to which occupational barriers can be avoided properly.

2.5 Big data / Machine learning models

Information system development process helps in managing the level of equivocality and uncertainty of project operations. It can assist project managers to improve management processes by arranging relationships between control systems and project performance. As followed by Sakka et al. (2016), through this process, project managers and project members can easily connect with each other and maintain the flow of execution within project management. Therefore, they are able to provide required results through their project management operations by enhancing project performance. Through the ISD process, project managers can make proper decisions and also increase performance rate in project assignment. Using this process, project members can identify their roles and responsibilities and they can give their potential performance to achieve their project objectives and goals. It can control every segment and activities of project operations through which project managers can easily increase project management operations. It can bring a stability and accuracy in service operations through which relationship between control system and performance management can be successfully organised. Bayesian approach helps in managing earned value management through which estimation process in project management can be organised successfully. Using this approach, project managers and members can identify required essentials for improving performance management to develop proactive management. It can assist in gathering specific knowledge through which accuracy and performance can be improved to manage complex situations. In the words of Caron et al. (2016), it can develop the structural way of arranging planning and other activities within a project assignment. Through this approach, project managers can easily make suitable decisions for making benefits for their project assignment. It can develop the way of analysing estimation processes through which internal and external operations can be organised properly. It can give opportunities for project managers to share their views and other activities to reach their destination in a given time frame. It can also develop the execution process through which stability and accuracy can be organised properly and also develop the way of managing every task within a project assignment without any hesitation. Build operate transfer model helps in providing a structure of project life span through which project managers can organise the sequence of project activities. It can allow project members to execute their tasks by following this model to gather specific requirements of a project and also maintain the structure of construction procurement. Thus, project managers can take appropriate actions by arranging all tasks and requirements and build a proper infrastructure to manage construction processes. As pro-

posed by Zhang et al. (2016), through this infrastructure, project members can execute their tasks by following a proper project plan through which period of every segment can be organised properly. In this manner, all operations in project operations can be organised and optimal solutions can be included within a project assignment to deliver possible outcomes towards their clients. Public private partnership strategy helps in allowing governments to provide financial support to maintain budget within project operations. Through this strategy, fundamental requirements and other project related operations can be maintained. It can allow in managing all analytical operations in a synchronised manner through which stability in the working environment can be executed effectively. However, due to lack of proper knowledge in project activities, project members and managers cannot understand the way of arranging sequence of project operations. As opined by Chou et al. (2016), it can decrease the confidence level of project members and decrease the progression rate for project operations. Government can use PPP strategy to develop service efficiency and manage functionalities to organise project expertise's to organise public works. It can assist in managing operational management due to which stability and accuracy in service operations can be executed appropriately. Thus, complexities in project operations and functionalities in public works can be maintained properly that can manage operations, preparations and construction processes. Through Artificial Neural Networks, predictions and detection operations can be organised successfully within a project management.

Analytical approach and resource allocation related operations can help in optimising location allocation problems and decision-making processes. Through the assistance of this approach, data driven operations can be developed and improve decision making processes. Thus, project managers and members can execute their technical and project operations and bring a positive atmosphere within project management. It can develop the way of execution and project progression due to which stability and accuracy in service management can be maintained properly. As proposed by Doolun et al. (2018), on the other hand, due to issues in hybrid evolution, project members and managers can understand the sequence through which effective results cannot be gathered properly. It can slow the development criteria and also slow the process of development process for project management operations. The project manager can make proper decisions in developing their existing data driven processes using analytical approach to improve performance quality. It can assist in improving quality and other operations without any issues in this current situation. Therefore, project members can understand project objectives and take proper steps in accomplishing their tasks. It can allow project members to deliver quality essentials depending on re-

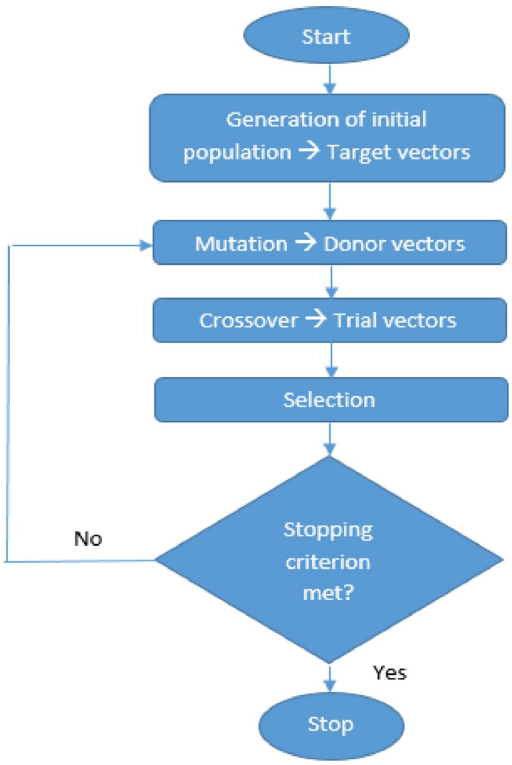

Figure 2.4: Flowchart of DE. (Source: As influenced by Doolun et al., 2018)

quirements to maintain quality in project management. Heuristic model for sustainable procurement helps in analysing change management operations through which project members and project managers can take proper steps in managing change management. Big data helps in managing sequential structure of supply chain operations through which logistics and procurement operations in project management can be developed. Through heuristic models, essential characteristics of big data can be developed through which optimal solutions can be executed properly. As opined by Kaur and Singh (2018), it can help in improving decision making processes due to which structure and execution process of project management operations can be maintained properly. It can develop the process of decision making through which structure and optimal solutions within project operations can be maintained properly. Due to lack of knowledge, project members and managers have failed to understand the way of managing all tasks due to which accuracy in service management cannot be maintained properly. Thus, project managers can analyse that situation and take proper steps in avoiding complexities and improve sequence of project activities. It can assist in dealing with any kind of situations through which stability and accuracy can be managed properly.

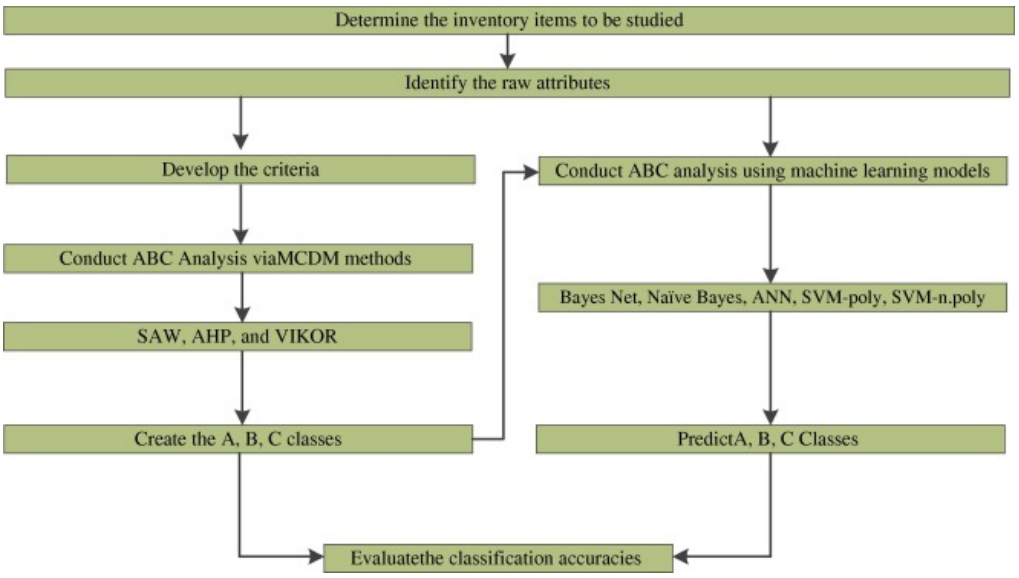

Figure 2.5: Decision analytic framework. (Source: As influenced by Kartal et al., 2016)

Decision analytic framework helps in presenting a structure of decision-making process through which project operations and other network related operations can be executed effectively. It can generate a way of managing sequences due to which stability in decision making process and detailed prediction performance can be achieved properly. As influenced by Kartal et al. (2016), it can assist in incorporating innovative ideas through which stability in performance management can be developed within project management operations. Through this framework, project managers and members can avoid complexities and other complex situations by reaching its destination.

2.6 Summary

From the above study, it has been understood big data plays an effective role in arranging sequences of every task and gathering specific information due to which accuracy and stability can be organised properly. It can assist in allowing project members to execute their tasks and also develop the way of managing all kinds of tasks depending on requirements. It can allow these members to make proper decisions through which sustainability and accuracy in service operations can be organised properly. Therefore, project managers and operators can achieve their goals and objectives through which sequence of every task in project operations can be achieved properly. Using several models and frame-

works, project management operations and project sequential tasks can be managed properly. It can assist in increasing quality and also develop the way of managing any kind of sequences without having any complex situations. Literature gap presents shortcomings of research study due to which opportunities for developing quality in research can be maintained properly. From this study, it has been identified that more theoretical knowledge about the role of big data on project management is required to be included in this study. As these articles do not have accurate information about theoretical overview regarding this topic. It can decrease the quality of this research study that can slow the process of managing information within this study. Moreover, lack of information regarding social mapping can decrease the structure of literature review within this research study. Many of the libraries and journals are not accessible due to which more information about big data analysis on project management cannot be accessed properly. It can reduce the practical knowledge from this study and also decrease efficiency in understanding skills. Therefore, in future, all gaps can be resolved and develop the quality of this research due to which stability and accuracy can be achieved properly.

Chapter 3

Research Methodology, Analysis and Findings

3.1 Research Methods

This study aims to analyze the impact of a suitable selection of procedures for information collection regarding big data in improvising management in companies. It identifies Qualitative research methods followed by instrumentation that aid in meeting research aims and objectives. Qualitative research is a theoretically driven enterprise, different from but complementary to quantitative research. It finds out about people's experiences, understands what is important for people. It complements quantitative research in particular by entering into the 'black box' of how social phenomena (including interviews and focus groups) are constituted in real-time. It is as much about social practices as about experience.

Project members are frequently obliged to circumvent (work around) rules. Qualitative methods are essential to discover, evaluate, and documenting how and why project members often could not follow the required technical and workflow procedures. It shows how organizations can function most effectively when their members bypass certain rules. Workarounds have characteristics of social problems - common practices that deviate from norms or rules. Engaging in them may risk harm or censure, but allows people to accomplish goals that are otherwise difficult or impossible to achieve. Ethnographic methods help identify faulty design assumptions, actual workflow realities, and belied over-reliance on faith in technology. Most workarounds work. However, they often obscure problems and possible solutions. We all use workarounds, actions that circumvent or temporarily 'fix' impediments to achieving a goal or reaching it more readily. Workarounds are ubiquitous because some prescribed rules or processes don't make sense, or are impossible or

harmful.

Workarounds are a frequent but unofficial aspect of work as well as much of life. From an institutional perspective (a business; society as a whole), the workarounds are deviance and are viewed as social problems (Best, 2017). The institutions and other individuals might ignore the problem, condemn it, punish it, or sometimes seek changes in the process leading to the workarounds. Punishing the individual is common, whether it is by financial penalties for violating procedure, jail for criminals, or institutionalization for those deemed insane. Some, however, informally applaud workarounds, as when a worker shares the password to a supply room with urgently needed equipment. While organizations view workarounds as individual deviance, workplace rules or technology limitations often necessitate or encourage them.

When the process often fails to work as designed, and medical staff use many ad hoc workarounds to get patients the medication they need (Patterson et al., 2002, Patterson et al., 2006).

Sociological methods - mostly ethnographic methods, but used in conjunction with quantitative data crunching - that are used to discover and evaluate the workarounds (Ching et al., 2014; Koppel, 2014: Koppel et al., 2008).

Ethnography permits the development and pursuit of unanticipated behavior and factors. The much-cited approach of grounded theory is based on ethnographic methods exactly because that methodology aims to build understanding where little exists (Bryant and Charmaz, Glaser and Strauss, 1967).

Both project managers and project members view workarounds as choices of individuals, but in fact they are a widely used set of practices found across organizations.

Workarounds are not usually the result of project members taking shortcuts but instead are mainly generated systematically by Project's faulty design assumptions, workflow realities inconsistent with design, organizational blindness, vendor hype, and the complexities of project.

Penalizing those engaged in workarounds is an ineffective way to prevent the behavior. The effective response is to fix the flawed designs and improve workflow rules and organizational policies (Holdem and Karsh, 2007; Karsh et al., 2006; Koppel et al., 2005; McDonald, 2006).

Actions of systems, institutions, and individuals have multiple consequences, some of which likely are not recognized by actors. Moreover, behaviors viewed as deviant can be beneficial (Gans, 1972).

Organizations usually dictate processes to which most people conform. When people do not conform, they sometimes displace efforts from the approved goals (Merton, 1968).

Observations can be used to learn the actual sequence of activities

leading to each 'Project workaround'. Project members comments will be useful to find the causes of workarounds.

If we have a record of the perceived causes of users' actions, these entries could supplement explanations that were verbalized or observed even if not explained by Project members.

Insights about workarounds can be interpreted by persons knowledge about the technology and work processes.

Workarounds are usually due to limitations of the technology, lack of staff knowledge, missing or unreadable data, or other factors.

- Technology-related (Software or hardware)
- Task-related (Protocols with which users are unfamiliar or believe slow their performance)
- Organizational (Organizational protocols and activities incompatible with safety)
- User-related (Project members decisions for special circumstances)
- Environmental (Location)

Multiplicity of relationships among causes and workarounds could have a large impact and be hard to eliminate. Each cause can lead to several workarounds. Some causes can be related to many workarounds and vice versa - a challenge for prioritizing improvement efforts.

Ethnography is suited to a project of discovery, analysis, and developing corrections. Ethnography is key to discovering and documenting workarounds and their causes, and to understand how they are generated in the work process.

Methods of researching and analysing gathered evidence in a research takes place by "philosophy, approach and designs". According to Bishop and Kuula-Luumi 2017) by applying apt "research methods", researchers are able to select procedures for collection of information with ease. This additionally helps in reflecting appropriate outcomes in study based on specified research topic. In case of this study, it has selected a "qualitative research method" for accumulation of data related to analysing effect of big data in managing projects in companies. In addition to this, it can be conveyed that research methods comprising "positivism philosophy" along with "descriptive design" and "deductive approach" have been applied. Apart from this, it can be elucidated that implementation of these methods has helped for researching on impact of big data analytics for managing project assignments in companies. According to Saunders et al. (2015), these methods have focused on

developing techniques of gathering data by enhancing facts and observation followed by making changes in existing theories. Thus, application of these ”research methods” has justified fulfilling objectives based on identification of challenges faced during project assignments in companies due to inefficient formation of data analytics.

3.2 Data collection and analysis

This study aims to analyse the value and contribution of big data, which is being better perceived by project members from varied fields of work in managing a project. This study interprets qualitative research method for a gathering of data followed by ”interviewing Project/Product managers, ML Engineers” in terms of meeting objectives of research appropriately.

Qualitative Interviews actively construct meaning in each other's talk. Interviews reveal evidence of the nature of the phenomena under investigation, including the contexts and situations in which they emerge, as well as insights into the cultural frames people use to make sense of these experiences and their social worlds.

Positivists seek to create the 'pure' interview - enacted in such a way that it comes as close as possible to providing a 'mirror reflection' of the reality that exists in the social world.

Naturalists suggest that unstructured, open-ended interviewing elicits 'authentic accounts of subjective experience'.

While this approach is seductive, a significant problem lies in the question of whether these 'authentic accounts' are actually, instead, the repetition of familiar cultural tales.

Finally, radical social constructionists suggest that no knowledge about a reality that is 'out there' in the social world can be obtained from the interview, because the interview is exclusively an interaction between the interviewer and interviewee in which both construct narrative versions of the social world.

The problem with looking at these narratives as representative of some 'truth', according to these scholars, is that they are context-specific, invented for the interactive context of the interview, and representative of nothing more or less.

There is a considerable difference between being skeptical about the bases of truth claims while carefully examining the grounds upon with these claims are founded and denying that truth - as a utilitarian and liberating orientation - exists at all. (Sanders 1995: 93, 97).

Qualitative interviews provide us access to social worlds, as evidence of both 'what happens' within them and of how individuals make sense of themselves, their experiences, and their place within these social worlds.

Interviewers need not resort to romanticism, or to identifying experience as authenticity, in order to utilize interviewees' accounts to provide authentic accounts of social worlds.

A traditional view of the ideal interview is that of a neutral conduit for excavating and conveying undistorted knowledge. Researchers recognize the interview as a meaningmaking conversation - a site of narrative production, equal with, not superseding, other forms of narrative production. Interviewing is unavoidably interactional and constructive. It actively produces knowledge in practice. All approaches for interviewing rest on images of subjects behind respondents and interviewers.

Conventional approaches envision the subject behind the respondent as essentially passive and the subject behind the interviewer as essentially disinterested. Images of subjects behind the interview participants have important implications for how the interview process is conducted and how interview data are constructed and appreciated. Interview participants concertedly engage in narrative work. This work takes place under the discernible auspices of narrative environments, which provide the conditions of possibility of intelligible and accountable interview narratives.

The versions of meaningful experience that emerge from interviews are constituted in the reflexive interplay of the hows and the whats of the process. The concept of the active interview casts interview 'bias' in a new light. All participants in an interview are implicated in the construction of narrative reality. They are involved in narrative production, not contamination. The guiding question should not be whether interview procedures contaminate data, but how the interview constructively generates the information it does. Because interview data are products of narrative practice, data analysis demands a rigorous sensitivity to both the hows and whats of the interview process.

The active interview is not a particular type of interview, to be distinguished from other forms of interviewing. All interviews are unavoidably active communicative enterprises. Even the standardized survey interview is active, because standardization actively structures the interviewer's input and restricts the respondent's range of responses. By calling attention to the constitutive narrative activity inherent in all forms of interviewing, we are pointing to the wide array of interviewing practices that construct experience rather than privileging any particular version or foundational model of the interview.

At the same time, by focusing on the active character of interview narratives, we are not saying that 'anything goes'. Put into place, every image of subjectivity spawns its own operating rules. The concept of the active interview derives from an ontologically warranted basis for interpreting the production, collection, and analysis of information in a partic-

ular way, and demands its own set of procedural and analytic guidelines. Hong et al. (2018) explained that application of ”qualitative analysis” help in managing analytical procedures followed by ”qualitative” ways for clear explanation. Information regarding various projects, internal processes have been collected by interviewing Project managers, Product managers and ML Engineers working in various companies. In this manner, this study has opted for ”qualitative research methods”, which has been significant in analysing effect of big data for managing project assignments in companies. Qualitative research methods help in gathering specific information regarding issues faced by companies to maintain project phases without using big data analytics. Moreover, this method helps in discussing authors' perspectives to analyse critical views on project management processes without big data analytics.

Collected information can be effectively analysed by ”qualitative” that determine to provide suitable outcomes. From this, it has been understood that incorporation of these methods for accumulation of evidence and as per accordance interpretation has successfully justified requirements in this study related to big data contribution for managing project assignments.

3.3 Ethical principles

Some classic terms in research ethics such as Consent, Confidentiality and Trust. Codes and Consent - refer to *informed* consent. This means that research subjects have the right to know that they are being researched, the right to be informed about the nature or purpose of the research, what to participate involves, benefits and risks, and the right to withdraw at any time as well as the terms of withdrawal. The tendency to archive qualitative data complicates informed consent because no one knows what and how data will be used for secondary analyses. We tend to think of research ethics as an issue between the researcher and his or her research participants, but it may also involve persons who appear in the participants' narratives. This raises the question of who our participants are, and of how to think of the scope of our research ethical responsibility or care.

The combination of informed consent and covert research serves as the ultimate research ethical problem though also with its own dilemmas (Punch 1994). Punch also reminded that the more deviant and secret the activity, the more subjects may fear disclosure and the bigger the responsibility for the researcher to protect them.

Confidentiality means we are obliged to protect each participant's identity and the location of the research. Qualitative researchers practice

this by using pseudonyms or by blurring personal identities, which makes anonymity one form of confidentiality However, anonymity by pseudonyms also carries connotations of gender, age, ethnicity, class, etc. where readers assumptions and the author's intentions may collide. This makes some prefer to use numbers. Another way is not to share all data with the reader. However, without the many and detailed events narrators employ in their storytelling, narratives and biographical research would suffer.

Yet we cannot always assume that participants want to be treated anonymously. To claim they do not know the implications could be correct, but it might also be to underestimate their wisdom and overestimate our own reflexive capacity (Nortje et al., 2019, Trimble and Fisher, 2006, Silverman, 2017). Anonymity means they lose ownership of their own stories and will not get credited for them, not even for their political activism when relevant. Sometimes researchers try to protect the identities of their participants when perceived to be members of a vulnerable category as in studies with social work clients, victims or illegal immigrants. If not assured of anonymity, storytellers may leave out important details of their stories. Kaiser (2009: 1632) refers to internal confidentiality or deductive disclosure when our rich descriptions make individuals or groups identifiable in our reports. Gerver (2013), based on her research on refugees, defends exceptions to blanket anonymity despite risks whereas Walford (2018) argues that the growth of social media and digital communication makes anonymity in ethnography impossible. To sum up, as Shaw (2010: 8) reminds us, ethical decisions are never decontextualized.

These issues also include harm and intrusion and the need to avoid common-sense assumptions of these and other issues such as privacy and 'sensitivity'. Privacy goes beyond the content of disclosure to imply even the audiences and circumstances involved. Also, there is no simple formula for sensitivity as in the link between qualitative research and emotional stress, and no bureaucratic documents can ever tell us exactly when or where these boundaries are. External control is not necessarily the answer to emergent challenges.

Trust refers to the relationship between the researcher and the participants, and to the researcher's responsibility not to 'spoil' the field for others in the sense that potential research subjects are reluctant to be studied (Ryen, 2004). In this way, trust also applies to the report or the discursive practices defining the standards for presenting both the researcher and the work as trustworthy (Fine, 1993). Trust is the classic key to good field relations and is a challenge constantly unfolding during the research process, though more so in ethnographic studies than in other kinds of fieldwork.

At times we come across delicate situations that involve hidden or

problematic information where someone may be harmed or put at risk (like crimes being planned), or where certain findings may be discomforting (like job evaluations, health or other information, Ryen, 2012) or even dangerous to some subjects (like some kinds of illegal activity). This calls for decisions on whether or not we should do such projects (alone or with someone else), if there are data we do not want to collect (we may turn off the recorder or the smart- phone, or move the conversation onto another track or be more explicit), or if we simply need to shut down the whole project. In some way or another, all these issues are linked.

Any model based on a pre-fixed, rather naive, ideal research subject (or rather, object) fails to capture what is special about qualitative research. With reference to the classic concern ”Can we trust them?’, questions of validity and research ethics such as confidence and trust become closely entangled, especially in the data collection phase.

The response in positivism is based on the assumption that language refers to an external reality ’out there’ that makes access to the assumed reservoir of stocked information crucial. By contrast, in the constructionist model, social reality is a more complex phenomenon where they examine how members produce recognizable forms that are treated as real or ’worlding to cite Gubrium and Holstein (1997: 42). In this model, the stories they get are produced with rather than by someone: they are contextually produced, designed for particular audiences, serve purposes locally produced and embedded in wider cultural contexts.

As researchers we have a responsibility to produce rigorous research. However, just as research which some see as dubious may provide us with data otherwise difficult to get (such as in covert research, or the fine balance between adequate and too much information when recruiting informants, see Fine, 1993), naive simplified assumptions about field relations pose another dilemma.

The bureaucratic procedures of research ethics, governance bodies tend to tilt towards a positivist epistemology with an audit culture and mechanical application of ethics, codes that are incompatible with most qualitative research. Constructionists argue that the social world is constructed or collaboratively accomplished. This makes codes and consent, confidentiality and trust more complex phenomena and informed by method and epistemology. Despite a tendency towards standardization, research ethics procedures and guidelines vary by global region, country and indigenous communities. Ethical clearance is mandatory prior to project start-up.

Ethics have been managed in this study by implementing ”Copyright, Designs and Patents Act 1988 (c. 48)” and ”Data Protection Act 2018 (c.12)” conveniently. Researcher has assured that data has been pre-

served in the laptop. Consent to record the interview has been taken prior to initiating the interview process from all the respondents. By such manner, ethics have been successfully maintained in this study, which equally focused over integrity of information.

3.4 Data collection method

3.4.1 Scope of the interview

The scope of the interviews is to analyzing the perspectives of project/product managers, ML Engineers regarding the usage of big data in project management. The interviews also tried to get know the way management of projects is done in corporate companies, their priorities, future goals etc. Through these information, positive and negative aspects related to this concept can be identified properly that can help other managers to understand the processes or techniques to use big data within their project management operations. This information also helps in identifying practical knowledge related to this topic through which structure and execution approaches can be analysed properly. It can develop ground knowledge and increase awareness concerning big data usage in project management operations. Issues and mitigation strategies can also be identified through this interview of qualitative method that can develop the quality of this research as well.

3.4.2 Qualitative method

Interview with Product Manager Dane Meten Weil
(`https://www.linkedin.com/in/greatestdane/`):

Dane is working as a Product Manager since the last 5 months, and as a Senior Software QA at GameChanger (`https://gc.com`) and completed his Project Management certification course recently. GameChanger is a mobile app-based software company owned by the parent company DickSportingGoods (`https://www.dickssportinggoods.com`). Dick's Sporting Goods retailer sells, sporting equipment and clothing and tennis rackets and bikes. They are the biggest retailer in the United States for sporting goods. Below is the summary of the Interview:

What would be your day-to-day work? How is it organized? What activities is it comprised of?

GameChanger is primarily a mobile app-based software company, we release new versions of app every single week on Mondays. The rhythms are Daily and Weekly. On a Daily basis - the Product manager of the company almost servers as a bit of a Scrum Master, they start out with team wide stand-up, as a cross departmental integration, we get everybody on call, get in line with accomplishments of prior day, what we got going on the current day, any blockers that may exist, rehash priorities for the current Sprint/release cycle.

After that I meet with other Managers on the team such as Engineering Manager, QA Manager, Lead designer and have a higher-level talk on how we are doing in the current sprint cycle? What work we need to prepare for upcoming sprint? While the rest of the Engineers Team are heads down working on the current sprints (stories, cards, tasks). We also have a Retro meeting at the end of the week for everyone including managers and reflect on the productivity for that week, what we could be doing better? what we are proud of? Things that we may want to implement in the next sprint to improve our productivity. These are the set of activities on a daily and weekly level.

At a monthly level, we have our QR reviews, monthly meeting to talk about where do we stand against the monthly/quarterly QRs we want to achieve. It's pretty much leading those meetings, leading that planning and then I manage our project board which we run in a Kanban style. We have a triage column of stories and tasks to complete. I work with the other leads on our team to triage new things that come in and those will feed through our project board. There's a 'backlog', from the backlog, things will go into a committed column. Once an engineer has completed something, he will grab something from the top of our committed column. It'll become 'in progress', that moves to 'needs verification', QA is the gatekeeper from our 'needs verification' column and then to complete itself. So, we operate in kind of Scrum and Kanban sort of Frameworks methodology.

Who are your target audience?

Our Market is a two-faced or a two-sided. We have our application technology geared towards amateur sports. So, when we are marketing and selling our technology, we look at it as, we have the staff, the coaches, the administrators of these sports leagues and teams, who tend to usually be the decision-makers on what technology their teams or leagues use.

So, we gear most of our marketing and sales efforts towards them. But we actually make our Revenue off of the parents' players and fans. So, we view them as the community of the team. We have the staff

and administrators and the team community and those are kind of the two sides, the community pays to get access to team content, stats, live video of the games happening of their kids and we give those tools to the coaches and the administrators for free.

So that's really our target audience and Market. We do also sell our batches of big data to professional sports teams because they use it to recruit and Scout high school and college level athletes. So, we do interact with them, but it's not a primary source of Revenue so we don't really focus on them, but we do have relationships with them and remain aware of what their needs might be as well.

So, the prime source of income is on the apps, the Amateur Sports site, where the coaches identify the players? And then host them on your app, right?

Yeah, so they'll create the roster, then create their team on the app and score all their games such that when parents can't make it to the game because they have work or let's say they're in the military and they're stationed overseas or their grandparents are sick or live in a different part of the country. This way they can watch and follow their children, relatives' action and coaches can use the statistical analysis that we automatically generate for the games based on their score keeping and they can get stats to improve team performance and guide them on team training activities.

How do you select these players? Is it based on the initial videos of their play or to select the top 5 or 10 in each game?

They'll use it like for example, let's say you're a coach of a soccer team in Italy for a bunch of 15-year-olds. You can go to one of the local clubs instead of back in the day, you used to have to score while your games are happening. You would have to write down on pen and paper. Okay, you know Xavier scored.

Today Philippe had an assist and we won nothing against Florence today and instead of doing that old technique, our app allows them to do it digitally. So, you're using your own players for your own team. So, I just enter in my roster of all the kids that are on my team and that's how I build my team. So, it's just a tool to replace traditional through laborers and manual sort of methods and all your real-life sort of on your club and it is like a continuous integration process, like an API kind of thing where it refreshes for every 1 day or 12 or something like that. It's a real-time API. So, parents can follow the games live, the live video streams. Again, there's like about a 20-second delay through the video service that we

use, they process our streams and so we integrate with their API and that sequence causes about a 20-second delay and when content is actually delivered to the fans that are consuming it, from that administrators who are at the field filming or scoring the game on their phones. But yeah, it's a pretty continuous delivery of content for them and any in our systems. Our development is a continuous integration sort of system where we are merging new code constantly. But again, we deliver new versions of the app only once a week. But when users are on it, the content is flowing in constantly.

What are the challenges you face? Sometimes there is delay in the streaming, sometimes one of the batches of data could be missing, some API wouldn't deliver.

Yeah, there are 2 big issues that we face. The first one is scaling and the issue with that is we have, an automated smart scaling system. So that way we're not utilizing too many servers and processors when we don't need to and the rhythms of sport make it such that our market is primarily in the US and Canada. And so overnight our systems will scale down because nobody's playing sports overnight and only during the summer games can happen sometimes in the morning, but during the school year, games typically will happen on a Wednesday afternoon, a Friday afternoon or over the weekend and it'll be hundreds of thousands of users and teams all of a sudden getting onto our app, logging in data, streaming new games and our system has to scale up. And sometimes things will fail, some of the boxes that we try and introduced our cluster will fail and we won't have enough processing power for the amount of data that we need to be moving through our system and delivering to our users and they'll start to see delays in performance. And so that's one issue. So, we're constantly working on our algorithms for scaling. Utilizing new technologies, trying new services in order to overcome them. We have a platform team focused on this initiative.

The other, even when we've scaled properly, sometimes corrupted data is formed on our apps. We have some users using old versions of our app. And the problem with that sometimes, it'll format data packets in antique ways or old ways and we have data transformation that is supposed to take this old format and convert it into our new data formats, that our database tables are looking for. And those Data Transformations sometimes fail as well. When that happens, we will have data that's not properly formatted or has a null value for some object that we're not expecting, and those can cause issues to users where it will appear as if their data is no longer sinking. They'll see it locally on their application, but our processors will fail to process that data and sink it back to other

devices. So, they'll score game and reach out to the fans and the fans report the data sync issues. And what's happening is, the data on the coach or the admin side is stuck in a queue on their app because it's in a format the transformer wasn't able to translate the data into the new format that it needs, and it couldn't properly sync up. And in those cases, we have tools to extract the data from their devices, manually process it, and manually make the necessary transformations. In those cases, we usually push the user to update their app, so formatting the data in the new way gets aligned better with our current data models. So that's another issue we have the Legacy utilization of our application.

For the scaling problem and utilization problem, which technologies are you currently using to address these problems? Do you use Hadoop or Spark etc?

I haven't heard of any utilization of spark. Our servers and all our instances are on AWS and we use red shift for one of our two databases. We have an old app that's hosted on MongoDB. Our new one is in red shift through Amazon. We use Bagel in order to introduce and manage our new instances, but they are all hosted on Amazon AWS, and we use a lot of their hosted tooling in order to manage stuff as well.

What are the legacy applications built to handle the data initially and later shifting it to the cloud?

We were on the cloud since the inception the company started about 10 years ago.

Are you happy with the AWS payment system based on what you get and what you give?

Yeah, I think they've done a good job to make more of their service pay-per-usage and that's allowed us to optimize our spend and align it closely with our usage. It's our biggest expense at the company and we're constantly trying to find ways to reduce that spend and be more efficient there. But any excess spending that's done, has been due to our own inefficiencies and not because of rigid payment structures or anything that we would attribute to Amazon and the AWS service. So, in general, yes, we are we're happy with their price points and their sort of payment structures.

Could you elaborate a little bit regarding batches of big data to professional sport teams?

For the first time we were approached by the New York Yankees, at some public Sports Tech events in New York City, got into conversations whether the technology we were using could be useful and realized that they could actually use some of the data that we collect and own, which is typically baseball score keeping data for the high school and college levels. We had never really considered this but they approached us and a say ”hey, could we buy a certain portion of your data and we’d like to probably do this annually” and so he said ”sure, let’s build a system where we can query and pull the exact data that you need to”. We have our data organized in such a way that we can we can isolate very particular segments of our data. For example, they wanted kids at the prep level who were from certain areas or for which we had a certain number of events for that. That way is a large enough sample size that their statistical performance was statistically significant. We sold that batch of data to them. And once after we did that, we realized there could be more opportunity. Maybe we should reach out to other professional teams. So, we talked to teams like the Toronto Blue Jays, the Boston Red Sox, a bunch of East Coast teams, the Baltimore Orioles and the Washington Nationals. Eventually Seattle Mariners at West and establish this relationship where we would provide them with scouting data upon request. Not every team that was back in about 2013 and since then not every team has come back every year asking for data, but regularly each year, at least a handful of teams will approach us and say hey we would like to purchase a batch of your data. So that way we can run it through our own systems and do our own analysis against it in order to recruit new players that they would like to draft.

How do you organize teams, do you face any conflicts while managing this work?

I would say one big challenge is our resource allocation, which is our ratios of Engineers to Designers and our ratios of Engineers to QA analysts. We have more engineering power than we do the others, we’ll have 1 QA on a team per average and one designer. And 6 or 7 engineers and that 6 to 1 or 7 to 1 ratio isn’t ideal. It means that we the engineers are almost always looking for new work to be done and it rushes our designer to put together new requirements and specs and prototypes and means that we have to sacrifice some of what we call our customer development process. Which is when we’re doing user interviews, market research, in order to decide what our next product is going to be, how we’re going to design our next product and when there isn’t enough time and the engineers are doing a lot of work, designers also have to get

involved with what we call Design QA. And so, a lot of their time is being is being used. Just queuing all this new work that our team is doing and they don't have time to think down the line.

And put proper documentation together for upcoming projects, the same goes for myself as product manager. Sometimes there's so much happening in the current Sprint that it feels like we aren't able to properly plan and anticipate upcoming Sprints. So that resource allocation is definitely a challenge.

Honestly, I would say early on during the pandemic. The remote circumstances were a challenge at first for communication and productivity. But over time we've implemented good systems and process where we are documenting all our work more closely in order to overcome the challenges of being remote. More of the conversation is formally happening in slack or formerly happening in Clubhouse, which is where we manage our project boards and any time, we have a conversation on a call. They're doing a better job of taking notes and summarizing it, and this actually has helped outside stakeholders who aren't involved in our team day-to-day, maintain their clarity as to what our statuses and where we are on projects. Because we've been remote, we've had to utilize these tools a lot better because you can't just turn and talk to somebody and get an update. And as a result, it allowed our outside stakeholders to get much better snapshots and Status checks on our work, because we're really using all these services and platforms the way that they're intended to. So, I hope that carries back into the office because that's also something that definitely challenges our team in context to what we're doing and where we're going. But the outside stakeholders, they lose that visibility sometimes and we don't document conversations that happen offline. So that was a challenge before, but I think the circumstances of the covid-19 pandemic has actually pushed us to alter our process to accommodate for the new circumstances and it's provided value, that I think will last beyond when we return to the office.

What are the strategies you use to manage Big Data related issues in project management?

We have a data team. We call it our DNA Team of data and analytics and they don't have a formal project or product manager. They have a senior lead on the team who kind of serves that role and another product manager at our company will sometimes sit in on their meetings to help them plan projects. Not myself, but another product manager who used to be a data engineer, so he's very well versed in their technology and their operations and they don't operate on the same sprint cycle and cadence that our product teams do so. Our product teams are building new tech-

nologies for the app, specifically new features and functionalities. And the data team is strictly managing this data optimizing the performance of our database, building our new data transformers and all that, and also managing our analytics tools, which we use Looker for a lot of our data analytics for data visualization and things like that. Their process is a little is different. It's not weekly. They don't have a release every single week that they have to finish, their deadlines are ad hoc. When they identify a new initiative that they want to do, a new project or if they want to change our database service to something else, they work for as long as they really need. They have a little more flexibility in their project timelines and that could change one day. But we feel that, their needs and the level of risk that we are willing to tolerate with their work means that we cannot pressure them to necessarily complete things within very fixed sort of time frames. We're owned by a public corporation and there are sometimes reports and analysis that we need to provide to our parent company and they'll need that by a certain time frame because it could be for some public report or you know, some public filing and in those cases will need to provide the necessary data points on demand and we're usually given a healthy heads up we have systems in place to make the queries necessary to give them that information. So, it's not too costly for us. But every once in a while, they want something new that we've never provided before and we have to you know, dissect our data and understand what the necessary queries would be, make sure the data is clean, scrub it for any duplicates or anything that we really don't think is should be in there.

How many types of games does your apps provide data for?

We're used to only be baseball and softball and now we're sport agnostic. Now, we do our app servers every sport.

Could you talk about Work breakdown structure or Gantt chart or PERT? Do you make use of these things or do you use any other traditional project management methods?

Yeah, when we are planning and trying to decide what we're going to work on or if something's worth doing, we don't have any fixed waterfall model or the agile model. On request we use lots of different tools that are common in this framework value proposition charts. We build work estimation charts, but nothing fixed and rigid that we always use. We left the design team do a lot of initial estimation and planning and though they're constantly trying new methods in order to do that. The head of our UX and product design department is in charge and they

love to try the latest and greatest and new methodologies. So, we almost use a value proposition chart to map out the frictions that our users are feeling with the given problem. And what are the values that we can provide them to try and just understand at its essence what the problem is but beyond that yeah, depending on if we're building a bigger feature or smaller feature will then leverage different sort of methods in order to build the requirements and specs for our new product or functionality that we're going to design. But yeah, we still operate as a start-up, so we don't have company-wide mandates for use. The VP of product doesn't demand that we provide a fixed justification for every decision we make. He's pretty loose on that and really just prioritises are we able to deliver and ship constantly and on-time and lets us do whatever our team see as most appropriate for whatever our current project is.

Would you like to add anything?

With regards to specifically around data and our project management. I certainly have a lot of feelings and thoughts about the Frameworks that we use, our methodologies, what fits our company now and what's going to be good for us down the line. But more specifically related to data. It's definitely became a more important tool for us to make decisions about what we're going to build. Product managers have a very close relationship with our DNA data team. And in fact, we're starting to have a data analyst which we're actually calling a business analyst injected into each of our product teams so that, we can we can make our decisions and analyse the market and analyse our opportunity in much more scientific ways than we previously did. Before we were using quantitative methods, we would survey our users and conduct research purchase, white paper sort of reports and work with customer development agencies that can provide data about our markets, but we can do even more and we've identified that. We have a ton of data ourselves. We can do a better job of mining to understand where our customer pain points are within our product and what features and areas of our app are really providing the most return on investment. So, we're starting to do that a lot more. We do a lot of event logging within our app in order to analyse the UX activity and analyse performance of our app in various methods and various manners. So, we're leveraging data more and more in order to plan and decide and build our roadmaps and analyse our performance and that includes even analysing our own team's productivity. The analytics tools within our project management platform that can visualize our productivity, rates of completion on cards.

We're looking at that more and more frequently. So, data is becoming pervasive in all aspects of our company and something that we're con-

stantly looking to integrate further into our daily, weekly, monthly sort of rhythms.

Interview with Machine Learning Engineer William Menten-Weil
(https://www.linkedin.com/in/william-menten-weil-944a15ab/)

Liam is working as an Applied ML – Software Engineer at JP Morgan Chase

How is your day-to-day work involved? Do you work with project managers related to Big Data? What is the workflow?

I work as a software engineer and kind of move closely with the data science team members to take their prototype code and any kind of ticket they've been working on and operationalize them for the production environments.

Yeah, so our group is like, you get assigned to a couple of project teams that are all kind of Big Data related. The ML kind of mandates, gives a little more clarity about what I was doing. One of the teams was high frequency trading where usage of Big Data is greatly leveraged for millisecond real-time.

Transaction logs create a lot of data. That was one project. The other one was natural language understanding through querying data bases, so not as much data but more kind of processing time for generation as well as for testing.

Regarding Natural Language Processing, do you access the data using any APIs or from the community of your customer base?

From our internal customer base, different analysts that work at different desks. So, like we're doing it with trade different regions. Bonds versus commodities for example.

How do you handle day-to-day business processes and does Big Data make an impact there

Maybe from a manager's side. They are starting to kind of aggregate data to help with the project management side of it. I know agile is big about velocity, tracking and kind of those metrics, but I don't know if it's really gotten to the scale where people are like kind of tuning their project management style based on signal or predictions on Signals around those metrics, but I could definitely see that being an avenue in the coming years.

Do you/project managers see any challenges in adapting to these technologies? Or does the technical team handle these tasks?

Our RPM is very technical, started as a data scientist. He kind of takes us more a hands-off approach when dealing with the kind of business needs. Usually, we schedule meetings with the customers to discuss the needs as a team. Between the lead developers, project manager, and the stakeholders on the customer side.

What issues have you/your team experienced while using these analytics? We know, these algorithms are not always accurate. Sometimes they can't detect outliers, exceptions.

I feel these would get implemented using an analytics suite, some other tool kit like Jira, look at Atlassian. Because, they're the ones who really capture all that data for a lot of teams. With their ticket systems we could very easily see people just relying on what those products provide.

They need to provide clear signalling that this is just like estimates and you know when it predicts that some team member is like underperforming only will be or something. Yeah, definitely get some other kind of a things in the mix other than that prediction. because yeah, AI models are a little biased.

What are the Machine Learning algorithms that are frequently used?

We use a lot of Deep Learning. Burt and kind of newer NLP and coding things. I'm trying to use GPT 3, although we got a license for it. For the for the neural network stuff, it's a lot of inception these days.

How does it scale with data while training the models? Do you use distributed scaling applications or is it cloud-based?

We use EMR for data processing, some teams use it for training depending on what type of models they're building. I've seen people use a custom bootstrappers with the MR Cluster to install RE, to help with training. We use a lot of large ec2 machines that we automate like bootstrapping of those and the training sessions with the data flow.

If you had a discussion with the project manager? What kind of discussion would it be like?

It's still very much based on the human factors. So, they see based on

performance, on what other people would say, what they can see during meetings, through one-on-ones. I think people are still pretty hesitant to rely on mechanical signals. In some of these other analytics that we've been using for the past 10 or 20 years there are some known issues. Like you said not everyone performs the same way, some tickets can create different velocities based on their complexity or lack of specificity. So, I think there could definitely be some more room in the industry to use data science to aid in the project management. But yeah, I don't think I've seen too much of it yet. So, and maybe your thesis is on to something.

When the data gets corrupted or lost or someone deletes the log files and it gets updated in one system but doesn't get updated in the other. Did you develop any knowledge base to handle these issues real time?

Yes. It is a lot of talk in the bank about protecting against like an adversarial attack. And detecting kind of anomalies in the data as they come in. So, whether they be like drifts or just random outliers or these like targeted attacks to change classifications.

How do these teams operate internally, the development team, QA team, the management? Do you follow an agile methodology or is it depending on the type of project you have?

Our group was very agile based. That being said it is always kind of a mix anywhere. I think most places at the bank moved to Agile, there could still be some real strict waterfall teams. It's just so highly partitioned and siloed that I only interacted with like a few other orgs within the bank.

Could you explain me any project from start to end? How it began and how it was delivered?

Yeah, I can definitely describe one. You'll be able to find it because a lot of these things like internal technologies or like underlying technologies too small. Like components of the bank, internal users leverage the technologies. For us it starts a lot with the consultation phase because there's a lot of potential products, projects and data that we could be interacting with. So, we want to try and figure out what provides the most impact with the least amount of work to kind of prioritize. So, we go through a lot of like talking with the users about what data they have on hand already, what they might be able to extract from their system and then we started talking about what problems they have. Because

knowing the data that we could use will help us inform what we could potentially employ out of machine learning to help solve problems. Some we take it from the data first perspective towards any pain points that might express and that's how we can kind of target what projects to do.

The pain points would be investing in better stocks?

So yeah, high frequency trading is a tough one for a pain point because it's a known need or desire to become more ML based as a high-frequency trading platform. The other ones are a little cleaner in terms of the taking the users pain and turning that into a project. One was, all of our databases have archaic terminology and when databases had like eight-character column name limits. So, nothing could be named, right? Yeah. It's just propagated over time. So, we have a tool that helps people find search for these weird terms and give them back a dictionary definition of what it might mean across the different systems and that was slow and there are some ancillary searches that you do off of that to find out what group of terms it could be in or like you can go from a group of terms to a set of specific terms. So, it was a natural classification task, the search task in and of itself like a pseudo-NLP task where we could optimize the search based on Nearest Neighbor and some other vectors. That was one pretty clean example of taking our customers pain point and the data we had on hand.

Some other teams were solving customer support, hotline identification tasks with audio. That was another interesting one. My other project was taking user questions from like a chatbot and interacting with their databases. It's a SQL query essentially in English and then tries to transcribe it to SQL.

In high frequency trading, do you have any proprietary trading application?

We're creating like low level models, kind of like building blocks for our algorithm that would then use those signals to who perform the actual trading like the reinforcement learning or regression and classification. I think the customer base is mostly internal. So, it's like how we provide liquidity when we're in big markets. We also had like an RL agent to optimize order executions over a non-brief period of time. So, if you have to dump, millions of shares you might have to do it over like a week of trading. And then that also spilled into a buyback algorithm, because that's almost the same thing as when to optimally buy.

What are some of the strategies to manage big data related issues?

We don't have any real kind of specific strategy that I can think of other than this yet kind of data first concept that I mentioned. I think that's like kind of key to the whole Consulting Focus that we do.

How do you like approach a client? How do they rely on your analytics engine? How do you promote it.

We deploy a lot of services. We create services out of a lot of these things in most cases. Sometimes you can provide modules that other people would use as like a reusable component, like customize more views depending on the requirement of the user so then they could put that into their system. Most part it's like restful services other than like that high frequency of model that gets deployed in a event-based Scala system.

Do you provide the support as well for these services?

Yeah, the general bug fixes we do. We're not on the hook for real time support. That would be some of the other teams. So much data becomes such a hassle.

Interview with Project Manager Sam Sterling (`https://www.linkedin.com/in/samsterling-cfa/`)

Sam is working as a Project Manager at FCLTGLOBAL, he is also CFA Level 3 certified.

FCLTGLOBAL is a think tank, research organization that focuses on Capital markets. Organizations as well as companies that are invested in by these investment organizations.

What do you work on?

It's more about promoting behaviours that our research has shown to improve the longterm performance of Investments and long-term performance of companies and to try to provide nudges or suggestions to investment organizations and companies that are more short-term in nature that produce lower return. It might be detrimental to society because of either lower returns or for other what we call externalities.

Who are your targeted customer base?

We have a wide range of members, some hedge funds, some traditional

asset managers like Pension funds as well as retail traditional asset managers such as Fidelity, Black Rock, State Street, some sovereign wealth funds managing the country's Reserves. And then several public corporations. So, you're probably familiar with Walmart or Dow Chemical is another one.

And so, they support us by providing a membership fee as their executive teams feel like our mission is good for society and then it also helps position them to their peers and to investors or other investors that they're doing something right.

Could you explain me a little bit about the client onboarding process?

When a new member signs on, we will set up a new meeting where we will introduce the main points of contact for us. So, whether that's myself or one of my colleagues on the FCLT side will be meeting with the respective contact on the new member, the new client's side, basically exchanging kind of our backgrounds as well as our role in the organization. And we also provide the clients with some documentation showing the lines of work that were involved in, occasionally the points of contact as the clients might not be familiar with us. So, we'll usually give them or offer them an organizational overview of who we are and what we do, so that they can mentally understand what we're going to be involved in, and when they see an email from us, what they can expect and then after those kinds of very intro conversation, will provide a rundown of what the research entails and some of the current projects that we have working to see if there's anything that they would like to get involved in.

When it comes to projects, how many different types of projects do you have?

At any given time, between 5 and 15 different projects that were working on internally, any or all or some combination of projects, our members are our fully allowed to participate in, where we try to position, participate in the project. To our clients it's not a requirement but an opportunity for them. Our clients can opt into them if they would like to, they don't have to participate in the whole five to ten projects at a time because that could be a big-time requirement.

Who manages these projects within the forum?

Some of them will be managed by me. Maybe between 2 or 3 because I'm still relatively new, but the majority of the projects are run by either the

research directors. So, they'll have between 3 and 5 projects each. Or other senior people that are involved in my department, which is membership. So, the client relationship management team.

Could you explain me about one or two projects, which you have handled so far since Inception to the end explaining the important phases?

I think one project which isn't necessarily an external project, but it's a more internal, trying to improve our own processes, which I've completed, was creating what we call an engagement dashboard. So, a software tool that shows how often and how deeply our clients are participating in our work. Like how many times are they emailing us? How many times are they coming to one of our research events or how many times is their CEO joining a CEO Roundtable that we organized and aggregating that data across the whole membership and seeing where there might be where there might be gaps where maybe some members are not sending us a lot of emails or replying emails or maybe there's a member that only attended one event last year that might be at risk. We would think that because they're not participating with us, maybe they're not interested in what we do or maybe there's a problem with how we're engaging them. Maybe they don't respond well to email but we keep sending them emails. We try to hold back by promoting our new products and as much available documentation and try to get in more touch with them to know their requirements better. We use we use that dashboard to say okay this client is kind of cold on us. What are ways that we can maybe try some new techniques about getting in touch with them or maybe their point of contact when we were doing new member on boarding, maybe that point of contact we didn't know left the organization or maybe they are not the right point of contacts because they don't really know that much about what we do.

How do you gather this data for this application?

It's all through Salesforce. We have licenses through Salesforce to use the platform and then for our events where we are convening our members for a research project or maybe convening our members CEOs for Roundtable discussion, we have an events program called CVENT, it's kind of like a CRM similar to Salesforce but strictly focused on events and it gathers all your event participation and attendance data. It has an API that sends that data to Salesforce. Also, we use Microsoft Outlook plugin for Salesforce. So that all the emails that are sent back and forth from members also make it into salesforce. So, it has all those touch

points as well.

Do you face any day-to-day challenges while managing this application?

Yeah, there's occasionally times when it's not working as expected and it might be that the dashboard reports. So, the dashboard is composed of like five different reports. Maybe I created the reports with incorrect requirements or for filtering and the data that is flowing in into these reports is not showing up because the filtering is incorrect, or maybe the report formatting is off which has caused errors in the past.

How do you identify these issues? Do you have a process or a knowledge-based for a new project member to identify and work around that issue?

We create a knowledge base document on these new dashboards and Salesforce and through our reports and Salesforce which lay out the filter criteria. So, it's usually fairly easy when it's just a filtering problem or maybe a data validation problem to root that out fairly quickly. Occasionally if it's a little bit more complex, which it can be and it's not data filtering or data validation then there might be a join problem between the data that's coming through from the external sources, from CVENT or Outlook, which we then engage. Well first the Salesforce knowledge base online. So, Salesforce has its own help documentation. And if it expands beyond that, we have a salesforce consultant who can also help advise on these external platforms.

Could you explain me a bit more about the data filtering problem?

Yeah, to give an example there was a report that had a filter on it that I thought when I created it was for looking at email activity from the past fiscal quarter and we knew that there is a problem because we started seeing that we only had about a hundred emails sent for all of our members. We have 60 members over the course of 1 quarter, that's way too few. So, we knew right away there's something wrong. It's not pulling in all the emails. Let's look at the filtering criteria and in the knowledge base document. I had copied the filter criteria directly from Salesforce and it showed rather than looking at the past fiscal quarter, it actually looked at a date range and the date range was set from September to December 2020 rather than looking at the last 90 days. So, it wasn't a relative filter. It was actually a concrete filter based on a set date range, which is obviously incorrect.

What about the data validation problem? Is it something to do with incorrect or corrupted data?

It usually does have to do with corrupted data coming from the APIs or the plugin in terms of 'from the plugin from Outlook' when we started to see fewer emails. Make it into our dashboard and we found that a couple of our team members had an outdated plugin that was sending Salesforce data in an old format and it wasn't logging all the data that was being sent in the email. We were missing the email addresses correctly, which help sync up which members or the clients that are associated with those emails. So that was one example of data validation issue. Another was from CVENT, it started sending different identifiers, which help to map which Salesforce contacts had attended an event and because it couldn't identify which IDs is associated with, it wasn't showing that our members that attended the events that they attended and so in that case, it wasn't an update. We actually needed to install a separate application completely for these events, which we had been notified about but we didn't realize that they were going to sunset the old application. So that was kind of one where we didn't listen to our vendor or we didn't understand how critical the issue was from the vendor and then we didn't apply it. We had to upgrade those apps for all the clients. So, this had to happen just for this event application and then all of the member accounts are within, those in salesforce also synced up with all the new application and CVENT and they were able to back fill all the data. So, all the old event data that was missing, they were able to apply new identifiers by joining the data. I think they also used a SQL based database to get the two to have the new identifiers which were needed.

Who handles database related tasks? Do you have a data engineer or database administrators?

Not really. It's technically me, but the SQL syntax is different from what I'm used to for Salesforce, so I usually rely on our Salesforce consultant. I do have some pre-set query which they provided me to give troubleshooting on basic steps. But if there's a new issue that comes up, I defer to our Salesforce consultant.

And for the regular SQL issues you manage it on your own?

So, there's nothing that we use or that we can touch that has regular SQL issues, CVENT uses SQL as its backbone, but we don't have database access to this event database. We also rely on a vendor for.

When a new project member joins and is not aware of these technical skills, like database management. What kind of challenges does he face?

I think anytime there's been a new team member that gets involved in the data. I will usually do an introductory training session about the usage of the applications relying on them for introductory tasks that just leverage the front end of the application. Once they are more comfortable with that and experience with the various tasks associated from there. They'll never really touch the database since even I don't really touch the database, we can export most of the data into Excel and leverage and look at it in an excel interface, which is a little bit more friendly.

Are there any like collaborations with big data for this project?

It's not critical to the project but the big data application that I would say we use a little bit is python using the Anaconda, jupyter notebook to just use really basic querying like "describe" to show averages, distribution, range, maxmin for various engagement.

I don't leverage the more advanced aspects of machine learning simply because there's not enough data points. I know that big data can be really descriptive of bigger data sets or it might be able to provide some insight for smaller ones, but only really working with between a 100 and 1000 records for a member. And often times less like 50 to 100 records is more common. And so usually using jupyter notebook and having "describe" function and mostly just using the pandas set for basic Excel manipulation is fine.

When do you expect an increase in your data, in terms of the number of members from different regions, when you want to leverage big data analytics to get customer insights while introducing a new project?

It certainly is something that our research team does with big data sets when they're exploring a new research project, often times with huge databases of financial information, trying to get trends as we get more members involved in our work, the membership team. The membership department will probably also look to data analysis tools to summarize information from our members and engage with us. Right now, with only 60 members, that's somewhat unnecessary. You can see from an Excel spreadsheet and maybe just a Salesforce dashboard where the trends lie, but as we get to you know 75-100 members and each of those mem-

bers may have 1000 or 2000 employees that are interacting with us. It will become a lot more. It will become a greater need. We're still building out that part of our business though, for the most part we work with only 5 or 10 employees at each of our members, but we do have lofty goals in the future. Take out a lot more employees that are members involved and just have a lot more clients members involved in our work, too.

Could you explain me about the other project you have handled?

Yeah, this one is definitely less of a data focus and more qualitative, philosophical, how investors approach to responsibilities. So, it isn't necessarily supported by data and we certainly are not analysing data sets in order to create an output. But what we are doing is making sure that we have participation from all of our members in the project. Which, is something that you can report on and have data points to, but like I said with only 60 members, I would not say that's a big data project necessarily.

To give you a better example of a project that I don't necessarily manage but I did participate in and had an associate analyst role in that had more to do with data, was one around the analysis of executive compensation for the companies in the S&P 500. And so that was one that we were using, we didn't have jupyter notebook capabilities installed in the office yet when we were working on that, we were relying on data sets from a service provider and using Excel as a tool via formula to note irregularities in the data that was provided by this data provider. My role was to validate whether the data had inconsistencies that might be either duplicative or that are following a pattern, that wasn't a real compensation data. That's one data project I got involved in where we saw that the same numbers were being reported over and over again for some variables of compensation. If you check that across what the financial statements were of the company, it looked like the data provider had either created a bad dataset or maybe analyse the same information over and over and repeated it for different years. Whom we rely for the data is also really important.

Could you explain me about the project management methods used?

We don't really follow a software development model in our research, it does have a methodology but it's more statistically based. What we're producing our whitepaper research reports that have often at the core a data element to show a support or rejection of a hypothesis. And in that methodology development, the research team will go through kind of what is to establish just thresholds that are appropriate that I think are

well accepted. To avoid things like sampling bias or if there's not enough actual evidence one way or the other to have a statistically significant outcome. That's established as a criterion in the project framing and so if a dataset or what we have available to look at a new project isn't suitable. We won't continue with the project. So, when the project is being framed the research team will first analyse if it's even supported or supportable.

And then we'll do the data analysis and present it initially to our members to get their take on why we might be seeing what we're observing and then out of that we try to create outcomes if what we're observing is actually a negative trend in the market, if it's bad for society, or maybe bad for just investing period.

Could you tell me the names of these projects?

Yeah, it's called the engagement dashboard.

The other project we're basically relying on individual experience. So, testimony about experiences with somewhat uncomfortable topic areas in the investment management industry. We're taking what's generally higher-level executives experience on certain things like for instance economic impact and how trade protectionism or tariffs might affect the way they manage a portfolio or how diverse whether racial or gender at Investment Management organizations has been approached and how as an institution they are able to address that and so it's really testimony and experiential description.

We take those interviews in there like very small group private conversations with a couple of our members and we use that to distil not just their answers about those specific topics. That's not necessarily what we're interested in. What we are interested in is why they felt that way about those topics.

If or when the organization members invoke emotion talk in the course of their decision-making, how do you identify that and do you take that into account as well?

Yeah. We definitely do. We'll take why their decisions are made and we often directly quote what they've said as far as why they analyse aspects of their work in a certain, why they are doing what they say at face value. The people that were interviewing are generally involved in like Enterprise strategy, so they are often reporting directly to CEO and they have a good understanding of the organization overall. They're not simply a manager or may be involved in a single department like operations or investment or finance or marketing. They are usually involved in kind of

a holistic view of the organization and are able to comment about the ”why” as much as the ”what” we’re talking about in terms of the topic that we’re concerned about.

Interview with Project Manager Bharat Maripi(`https://www.linkedin.com/in/bharatmaripi/`)

Bharat is working as a Project Manager and Business analyst at Amaris, he also has many certifications in the field of Project Management.

What is your day-to-day work? What is it comprised of, what activities do you engage in? Who are your stake holders?

Currently I am handling three projects. So, my day-to-day work depends upon the project progress. For example, now my projects are in execution phase, so my stakeholders are my developers and also my business project managers or business stakeholders who use this product then I’m going to deploy it into production. So, I’m going to talk to them and organize weekly meetings to understand the project implementation progress. If we have any roadblocks while execution, the developers will get back to me and I will try to figure out a way to can solve the problem with the help of end users, try to understand optimal solution. One more project is in planning phase. In the planning phase I will try to understand how many stakeholders are needed and which stakeholders are needed at what time and who are the stakeholders to be informed. Here I’m going to prepare a stakeholder mapping matrix and try to understand what is responsible, who are accountable and considerable and informed and depending upon it I’m going to understand, my number of man hours it is required and break down the entire scope into chunks called work breakdown structures. I’m going to prepare a path for that work breakdown structure and then I’m going to get a buy-in from the teams who supposed to deliver those work breakdown structures. These are the set of tasks that usually happen in the planning phase with resource planning and communication planning and also budget planning and scope planning. One of my other projects is right now in the initiation phase which means the project is just started, but we did not even get a time to understand that project in a detailed way. During this time period I basically contact the end users and also key stakeholders from the customer point of view who uses this product, try to align with them. How is the product working now and why this product they want to develop in a better way? What are the pain points they’re facing and we will try to prioritize the project in this way and then prepare a work breakdown structure kind of a chart and we are going to move this to the planning

phase.

Do you follow an agile methodology?

We follow a predictable life cycle methodology, also called the waterfall approach. But one of the projects we are trying to move to the Agile since I work for a very big organization, there is a project management office and we usually follow the process, templates recommended by them.

Right now, Agile is dedicated to query few projects and not handling those projects.

Have you handled any agile projects in the past?

Yeah, I have handled 2 Agile projects for four months. I am also an Agile practitioner too for those 2 projects and we followed the scrum methodology.

Could you explain me like little bit about this methodology?

Basically, the scrum is part of the Agile methodology where most of the IT companies prefer to use, and my product at that time is a software-oriented product. So, we have a product owner for this project, product owner brings up the use cases, product backlog. We discuss with the developers to make an estimation for those specific use cases. We then have a Sprint backlog meeting together with the agile coach (me) or a scrum master, product owner, development team. We try to understand as many use cases, user stories we can tackle and we are going to pick those user stories for the Sprint. Sprint is 3 weeks of time period where the developers will develop the product, the product owner will validate it and then we are going to talk about deployment in the production. This is how the Sprinter life cycle usually plans in the scrum. My project will be two weeks. We will allow the developer to organize their work and finish those user stories on a day-to-day basis. So, I contact the developers regarding status of yesterday, what they're going to do today, if they have any roadblocks so that I can as an Agile Coach or a scrum master I can help them to remove it. After two weeks, I used to organize a sprint review meeting with my product owner, check the user stories whether they are aligned with the product of his vision and or not. If yes, then we are going to take it as an acceptance criterion. Otherwise, we realign with the product owner and understand where the miscommunication happened. After the Sprint review meeting, I used to set up a Sprint retrospective meeting. During that meeting me and product development team and Engineers team try to understand how we can do

these things in a better way taking the feedback from every individual, try to improve the process. After Sprint retrospective meeting, we again go to the product backlog and we try to tackle more user stories. This is how the cycle goes off every two weeks.

Do the vendors terminate the projects?

In these kinds of projects, we can complete only when the product backlog is empty. Realistically speaking when we are using Agile, we really don't know what is the end product because this end product is continuously evolving according to the business needs or according to the market needs or customer needs or it is going to evolve in a different way. We really don't know what we are expecting as an end goal. So, we don't prioritize or prepare for a handover. We continuously take up the user stories. We develop and the cycle goes on and on one fine day, we realize this product has moved to a mature phase and there is nothing more to develop and will focus on other product improvements.

What are these products and who is your customer base?

One of the products is a networking product like the Telecommunication. We are developing 5G technology. So, the consumers are everyone who want to use the tech 5G technology. The other project where I used Agile and subscription-oriented product. It's like a Netflix subscription every month. We are going to detect some money from you and we are going to deliver you the product we have ensured you. Using this strategy, we can improve subscription.

And the products could be anything in these subscriptions?

Yeah, the product could be anything, it could be a coffee, it could be machines, it could be milk, it could be Netflix.

How do you reach the customer base before starting a new product?

We call it as a product market fit. As a project manager, I personally don't go to the customer and ask. There is a front face team with business managers or product managers or product owners. They are responsible to understand the customer insights. But in some cases, a project manager might also take the responsibility of ownership, trying to understand that this idea is going to yield a better result. So, in this case, I'll try to set up some meetings and understand which other stakeholders

are impacted by this product, how they're impacting and what exactly is the problem they're facing, and what are the potential Solutions we can route of it, and how the solutions can improve, whether these solutions are aligned with the company's vision and mission, will the solutions create a good impact on the consumer segment. If yes, we allocate some budget to the project and implement it.

Every project has a different priority. We have limited resources. I can't take every project into the priority. We have to understand what is the high-value project? If not, we do a cost-benefit analysis after discussing with the business managers, identifying the solution whether this benefit we are going to reach out or not, depending upon it we go to the planning phase.

How often do you interact with developers? What technologies do they use?

It is not mandatory to understand the technical aspects of a project. But as a project manager, I personally understand always to be empathetic with everyone, to understand the team's morale. I try to understand which technology they're using; how difficult it is to implement certain use cases or user stories. I won't ask them to implement in my own way. Developers have the full freedom. My main concern is, if they have any problem, I'll try to understand from their explanation and try to help them with a solution. We can achieve this only when we know about the product, the technology in a better way. You need not be a developer for that, but you should have the technical edge on this technology or product you are using. This is the reason I always invest my time to understand the technology and the challenges they're facing and if they're moving from one technology to another technology this migration happens, I pay more attention to that as well.

Do you use the cloud or distributed computing services provided by vendors in the market or do you manage it on your own?

We use Oracle cloud services to manage our databases. I was part of the data migration from our physically located databases to a cloud database using the Oracle team. This is part of the cloud migration projects we have handled. We also have certain services that are in-house infrastructure only but those service providers are external. These services are part of our architecture. We have a complete autonomous control on them. But if any issue occurs, we usually raise a ticket to the application provider to understand whether this is an internal issue or an application issue.

When you were migrating from your infrastructure to Cloud, what kind of issues or challenges did you face? Also, in general in project management, what are the different types of workarounds?

We do risk management planning to understand the potential risks that a project might face and the probability of it happening, it's impact. We measure the impact and assessment of every probable risk that project might face and for every impact we also think about the mitigation plan, whether some risks we have to accept and we have to move on, some risks we can migrate and some other risk we can mitigate using some workarounds and some of the risk we have to handle to a full complete solution. We have to stop our other timelines and we have to focus on this risk. So, it depends upon the kind of risk we encountered during a particular project progress time, execution phase and accordingly we are going to follow. If there is an unknown risk, we have a contingency reserve which means our resources and our budget allocated for unknown events if it happens.

Possibility of 10% or 20% if it happens, we can utilize those time to focus on them and our research teams, solution teams, business teams will focus on taking their own time, which does not impact the project in a worse way. But yeah, our timelines can be extended, but it is okay because those things are already planned. Let's say if there is any Force Stop or any other natural events happened, we need to prepare a business continuity plan like moving from one location to another location, during that we have to communicate to the Senior Management, will stop our project and we'll focus on business continuity. So, this entirely depends upon the issue and risk that we have encountered.

Did you face any of these challenges during the covid-19 pandemic? What is the impact of digital Technologies or AI or big data in general? What is your perception on that?

In my projects, we always work remotely, our infrastructure is completely built in such a way that we believed on remote working and also smart working methodologies. So, for me, there is almost a minor impact, but certainly I would say for products which are part of the factory buildings, those factories have shut down. The team output and everything seems to be fine. Only thing is morally that people are scared and they might be having a hard time. We understand that certain tasks will slow pace and productivity might lose a bit because people are going to be sick or there are certain things that we have to understand in this phase. I personally am more empathetic understanding my own team, not just focusing on

the project progress, but also understanding personally and professionally to help them moving forward. Regarding artificial intelligence, big data on the project management, I never worked on that but I have some views on it as I could see the progress of being migrating the data to a cloud and also there are certain projects, we are doing on a data governance and data catalogue management to utilize the decision-making. So, there is a huge amount of data that companies are generating. We have a huge customer base, but we could not able to utilize those customers in a better way by helping them to make a good decision that can improve the business value. This is a bit lacking in many companies and including my company. Right now, there is a team focussed on data cleansing, data catalogue management and governance to manage the data in a more efficient way not to omit the data through utilize the data in chunks. We are also establishing a data science team, their main goal is to focus on predictive analytics, to help the inventory management, maintenance of the machines, to help the customers to get a bright product at the right time with the right scope. So, there is a lot of decision-making process that we can leverage upon the data right now. The companies are progressing there, but this is not so matured. This is for us is in an inception stage and we really call it a luxury, because our operations everything is completely not dependent on the data, but we are using the data to make an impactful change in our organization. So, I would definitely say it has a huge scope in the project management.

How do you manage your data? To track the activities of your project members or if you want to go through a big file of the customers, how do you use access this? Do you use traditional Excel or like Salesforce for Outlook?

In order to manage the resources, we are using a tool called ServiceNow. This is a service management tool where we log our man hours, will understand the calendar of every resources that were part of this project to manage them closely. As a project manager, sometimes some people might not report directly to me, they might report to the functional managers. But the project is my responsibility to manage their time period or to manage their man hours, so we have a tool like service manager. Coming to the data, when I'm going to look for any issues or something. We use the Excel sheet. So far it is sufficing for us to focus on our work, as we are not a completely data-oriented company. We don't use data to make any decisions. We use a small data, but in my past, I used Tableau to create some charts out of data for good visuals and reports and try to understand pattern from the data.

How important is to possess the knowledge of python or distributed databases like Hadoop or Spark? Will there be a specific data science team that handles it in the future? If you get a project where a project manager needs a bit of technical understanding as well to work, do you think it works in that way in the future?

I personally think that python or Hadoop have a huge potential. Being a project manager, the goal is to deliver something on time on quality within the budget. These are the main responsibilities of being a project manager for a very long time period and now this is going to continue with certain changes for sure. I personally believe project manager should also focus on business impact not only delivering on time on quality on budget. But also, how to improve the business value. So even the project management is shifting towards this direction and it is shifting to also in some companies. The project manager is not only a planning guy or on quality person. In my job, I'm a business analyst also, I analyse the business, try to think how I can improve the business value and depends upon this I'm going to plan what is the high business value to the business value so that my end goal is to help the business to reach their best goals. Hence, definitely they should leverage their skills on python Hadoop not in a detail way to understand how to develop those products, but try to understand the potential of python. For example, how we can use python as a technology to improve their business and what could be the potential threats or what could be the potential challenge they should face when developing this project as a project manager. It is important for me to have the Vision on a high-level of taking any product or any idea or any technology into my pool.

If there is a communication gap between the developers and me and this gap will continuously increase if we are not giving enough to know them and understand their pain points. This can help me only when i want to know them in a better, when I know a bit about their technology, the potential challenges they might face, be understandable and it's a continuous learning process.

Is ticketing application service now, an annual subscription or how does it work?

This is a corporate decision making. This is completely depending upon project management office because they are the one who recommend us to use these tools. I'm sure that this is not going to be an annual or biannual subscription, it is going to be a decade because we usually make a decision considering at least five years of vision.

Regarding the Oracle cloud services, are you happy with the support they provide?

Right now, we are having some issues. Eventually we are going to stabilize. We initially got so many errors and works because most of the things are new for both the team members and mentors. It eventually created some regression issues during the implementation phase and our business has been impacted a bit. But this is quite expected for us. We have planned this kind of risk into the consideration, but now it has been stabilized and if we have any issues, we have a different service level agreement (SLAs), depending upon the issue's severity, we are going to receive the support from Oracle.

And are there any incidents in the past where some of the data got corrupted or lost?

There is no such incident of data corruption. But there is an incident that we could not retrieve the data. Some data could not be fetched in the way that we expected. It is due to some settings inside the database some configurations. Because when we are moving to Cloud there could be some changes that we have to do in the databases. Those changes were missing, but eventually we resolved it. We have Data Integrity; every data has been present.

Is this handled by Oracle as part of SLA or your internal team manages it?

We don't. We have a database team. They were working along with the Oracle team.

Would you like to add anything more?

You're going into a very interesting area where big data and project management are very essential in terms of making any successful impacts in the business, because project management seems to be a middle management work, but it is very essential for the company's growth. You have to deliver certain things to enjoy your top line and bottom-line improvements. Like revenues or profits, project management is the key pillar for that. You have to understand, execute in the right time with the right quality on the right pace. If anything is not happening. We are going to receive a strong feedback from the customers and we are going to lose the loyal customers. That's where the company says acquired from very long time.

Right now, decision making has become very qualitative. It should be rather Quantitative. For many years companies are thinking that they know everything about their customers, but when it comes to really seeing the data, they were surprised and the Customers Behaviours are changing very fast and keeping the data in their pool is not only sufficient. We have to make a meaningful sense out of the data. That's where Big Data oriented technologies can take a very great step in improving the company's one step ahead of their competitors or at least one step towards the future and this is very key and necessary aspect that it is going to be in near future.

3.5 Findings and Analysis

First interview reveals the usage of Big data in the Sports and entertainment industry. In almost all the interviews the respondents confirm most of the times Project Manager serves also as a Scrum Master and possess little bit of technical skills. Game Changer offers business through big data analytics providing insights such as Real-time stats from the digital game records (batches of Big data), Videos of games to Professional Sports Teams as well as to the community. Underutilization of the Servers, over utilization of the systems and the inability to handle too much data on big game days, data corruption, data synching issues are some of the issues faced on a day-to-day basis by GameChanger while using Big data analytics. AWS cloud service (Pay per usage) and utilization of relational and NoSQL databases such as MongoDB confirms the usage of both Structured and Unstructured data. While AWS service was the biggest expense for GameChanger, they are happy with the pricing structures. Resource allocation challenge in-between Engineers, QA, Designers, the possible workarounds designers may look for during tight deadlines. Covid-19 and remote working was also a challenge most of the organisations trying to deal with, it impacted productivity. But increased usage of Digital tools is bridging this gap and also documenting most of the work activities that are useful for future reference. Agile or Waterfall depending on the project, usage of Value Proposition charts. Big Data is a very important tool for decision making, increased collaboration with the Product managers and the data team.

Interview with William informs about the usage of Big data in High frequency trading, Natural Language Processing at JP Morgan. Projects handled in Agile manner aggregating data with project management, usage of Jira/Atlassian to handle tickets, AI models could be a bit biased sometimes, usage of Deep Learning (Neural networks), EMR for data processing, AWS EC2 instances. Project management is still very much

based on human factors, performance, opinions and impressions, big data analytics have a lot of scope to influence the decision-making soon. Preparedness towards handling adversial attacks, data anomalies indicate the stream lined processes linked with Big data models. Projects usually start with a consultation phase, prioritization, usage of Machine Learning to extract valuable information from the data, search optimisation based on Nearest Neighbour and other vectors, hotline identification for support, Chatbot interaction model, Reinforcement learning models, deploying services, provide support for bug fixes.

Promoting behaviours for long-term investments at FCLTGLOBAL with a wide range of customers such as Hedge funds, banks, public corporations. Consulting and engaging clients in the projects, managed by project managers or research directors by using digital tools such as Engagement dashboard, Salesforce. Issues like incorrect filtering, errors while formatting, data validation problem, synching issues between CVENT and Salesforce, were dealt through dashboard alerts and knowledge base. It is evident although Sam being a Project manager had to deal with SQL related operations and the data related issues sometimes and hence possess a bit of technical skills for the job. Sam confirms the usage of applications like Python, Jupyter notebook for basic data exploration but never dealt with Big data as there are not enough data points yet. Qualitative analysis is also taken into account towards investors. Statistical analysis is taken into account towards project management methods.

Project Manager Bharat mentioned day-to-day work depends upon projects progress, with a series of interactions of work among developers, stakeholders, project managers dealing with the man hours, work breakdown structures, resource planning, pain points. Using waterfall model, and sometimes Agile methodology with use cases for each Sprint. Projects usually terminate once they reach a mature status. Amaris targets its customers based on the product market fit. They use Oracle cloud services to manage their infrastructure. Risk management planning is prior done before initialising any tasks, projects, events and allocated a contingency reserve fund to handle unforeseen events. Amaris was already well versed with smart and remote working methodologies. But Covid-19 did impact production of products such as in factories. Still in early stages of adaption towards Big Data for Inventory management and maintenance of machines. Amaris uses ServiceNow to log man hours, resource planning. And some Excel and Tableau to generate some reports out of data. Bharat feels the responsibilities of a Project manager will include some changes and adaption to some technologies to improve business value.

From the above study, it has been found that big data analytics pro-

grams provide a structural way of arranging project management operations with the help of several techniques and methods. From the interviews, it has been found that most of the project members have agreed to use big data analytics programs for improving project management operations.

Most of the respondents have also agreed that machine learning and big data models help in enhancing business value and processing structure for project management operations. It can bring stability and accuracy through which project managers can easily maintain the effectiveness and gather specific results with the help of big data analytics programs as well. Implemented interview session to different managers, engineers have helped to understand that use of Waterfall, Agile models and work breakdown structures, Salesforce, ServiceNow, Machine Learning models are suitable for presenting changes in form of advancement in project management operations. This has helped to understand that taking support of such type of tools, techniques, project management operations receive effective outcome and accordingly able to perform well in real life projects. Additionally, application of data driven analytics aid in improving functions in project management operations so that it can operate well in project management activities. Thus, it guides for developing changes in operations of project management by the help of effective collaboration of criteria in terms of maintaining changes effectively.

On the other hand, project managers have also agreed that big data analytics is beneficial for their project management operations in terms of scheduling and costing operations. It can help them to maintain the flow and operational tasks through which issues and complexities can be avoided properly. However, one of the managers has explained that agile methodology is better than big data analytics due to lack of resources and enough data points. From qualitative analysis, it has been found that most of the managers have presented that big data analytics programs can reduce complexities and allow in making proper structure in terms of cost and scheduling operation.

It has been understood that the proposition of big data analytics tends to improve relations that exist between control systems and projects. However, it also helped in analysing that incorporation of big data analytics would be helpful for managing changes in project operations and yet project members need to actively participate so that during emergencies, possible mitigation strategies can be taken and help to resolve effectively.

It can help in understanding positive and negative aspects of using big data analytical operations. It can help in managing all segmentations depending on project management activities through which processing

and progression structure can be organised properly. Moreover, qualitative analysis helps in analysing the authors' perspectives to identify the usage of models and methods through which clarity in data management processes can be understood properly. It can develop effectiveness and also maintain the vulnerabilities in project management operations without any issues.

Chapter 4

Conclusion and Recommendations

4.1 Conclusion

Usage of big data methods towards contribution to project management has been entirely discussed and examined from collected external resources in this study. From a research background, it has been understood that accuracy of information related to the cost of marketing products is well analysed by the support of big data. However, this study proceeded further to analyse the impact of big data on project management in both positive and negative ways. The initial chapter has helped in implementing objectives so that further steps can be evaluated by research in this study. Apart from this, project management functions improve and optimal performances are achieved by regulation of big data analytics in operations. This study has focused on rational issues that emerge during functioning performance management. It also examined complexities that take place due to big data programs and lead to affect functioning of project management negatively. However, it has been identified that usage of big data analytics reduces expenditure in project operations along with enabling to incorporate new innovative techniques of working. This leads to ascending revenue to project operations and aid managers to increase working efficiency so that objectives can be met suitably.

The literature review has been interpreted with a conceptual framework that helps to present an overview related to significance of big data in project management. From this study, it has guided in understanding the impact of big data on the improvement of functions such as manufacturing, supply chain followed by machine learning models. In other words, it has conveyed that utilisation of big data procedures enhance structure for achieving sequential outcomes, due to which operations of project management improves. Additionally, the cyber physical system

improves with the help of operations of project management. Communication and sensor technologies enhance by applying big data accuracy along with the internet of things. From this study, it has helped in acquiring knowledge that supply chain management develops accordingly by a transformation of technologies that can be incorporated by big data analytics. Crucial aspects have been understood well as this study explained elaborately with incorporation of models and theories. Besides, this study has helped in analysing arranging natures related to improvement of functions that facilitate dynamic opportunities for project management to sustain. Data based on decision making process improved by analysing key requirements, which can be solely feasible by applying big data operations. Relevance of big data operations takes place by the help of understating skills and interaction followed by effective strategies and progress changes in activities of project management. However, it has been analysed that its application of P2P commutation is one of significant features that helps to identify issues and accordingly addresses issues so that performance improves simultaneously.

The Methodology chapter has collected information by applying qualitative research methods to meet objectives and aims of this research. In this study, involvement of ”3 Project/Product managers and Machine Learning enigineer” have guided in understanding the handling of projects, their challenges, issues and the increasing usage of Big data models towards decision making. Ethical principles are maintained by asking permission to record the interviews for reference purposes.

Project managers have agreed that use of big data leads to a decreased set of expenditure budgeting in projects as per convenience. Although the cloud services make up most of the budget expenses in a company, they are seen as investments in terms of ability to leverage data for decision making, selling batches of Big data and understanding customer insights. While some companies are still to adapt Big data technologies, this is a field all the respondents have agreed to revolutionize the business operations.

4.2 Recommendations

Project managers, Business analysts are expected to make appropriate decisions in incorporating appropriate strategies to enhance performance using big data analytics. Through big data, we can store equivalent information for every segment of project management operation. Using a big data process, project members can store data in an organised manner that can allow in executing accurate information depending on requirements. It can give structured results through which issues

and complexities related to progression rate can be decreased properly. Therefore, they can develop their effectiveness in project management operations that can enhance quality management processes effectively.

Integrating big data analytics

Big data analytics help in developing explicit knowledge, infrastructure, human inventory and financial operations within a project management. Through this process, project managers can identify requirements of every stage and provide accurate information towards their team members. Therefore, they can operate their tasks depending on the requirements to achieve goals and objectives without any challenges. Through big data analytics, performance and skills of project members can be developed through which they can utilise their skills to maintain the balance in planning and execution processes within project management. It can decrease issues regarding task and scope management issues due to which effectiveness in performance quality can be developed within project management.

Installing software for resource management

Installing software for resource management helps in developing quality in resources regarding project management phases. Through this strategy, project managers can easily provide required resources for their project members due to which quality and clarity in service management can be maintained properly. It can develop the procurement processes through which structure and execution related processes can be executed successfully. It can help in generating accurate results due to which sequence and sustainability in project management performance can be developed.

Training and development strategy

Training and development strategy allows project team members to maintain the balance in performance and skill management processes for project management. Through this strategy, project team members can understand their roles and responsibilities that can give the opportunity for developing performance rate. It can allow in generating appropriate results due to which sequence and development processes can be executed successfully. It can increase quality management processes due to which sustainability and sequence in performance management can

be developed as per requirements. It can increase internal and external knowledge for project members that can give confidence for them to sustain their stability in project management performance.

Data mining within project management

Data mining helps in storing accurate information regarding project management due to which quality in data processing can be developed. Through data mining, project members can operate their sequence regarding every stage of project management. It can decrease complexities in schedule and cost management processes due to which project members can deliver appropriate results towards their clients. It can develop the process of data mining and storage operations within project management operations. In this way, achievement criteria for project management processes can be developed due to which sequence and sustainability can be organised properly. Project members can store their valid information due to which sequence and structure using big data models can be enhanced within project management operations.

4.3 Linking with Objectives and Research questions

Project activities are handles on a daily and weekly basis based on the Sprint cycles using some dashboards and frameworks, rosters, player selection based on real time data points, optimising database performance, identifying product market reach. Using big data for high frequency trading, leveraging natural language processing for information retrieval. And relying on tools like Jira, ServiceNow, Salesforce, Atlassian to track and manage project activities. Big data pipelines are implemented through cloud services such as AWS EC2 instances, Oracle cloud services. Attribute classification task within the databases through algorithms like Nearest Neighbours, Chatbot interaction for providing basic support of the product, reinforcement learning to improve accuracy of the data models within a project. Project reports are generated through analysing data in Excel, Jupyter notebook using Python, Visualization tools such as Tableau.

Under/over utilisation of systems, data corruption, data synching issues are some challenges to deal with in projects when using big data analytics. In projects such as high frequency trading at JP Morgan, AI models can sometimes be biased and prone to outliers.

Strategies to manage big data related issues include adhoc workarounds

such as restarting services periodically, scaling algorithms, manual transformation of data when data transformations fail, tracking issues and documenting through tools such as Jira, Salesforce, manually configuring database setting sometimes.

4.4 Future Scope

This study provides information related to manufacturing, supply chain, project management activities and role of big data models. Through this information, a project manager can easily arrange a proper sequence for managing all the activities due to which achievable criteria can be organised properly. It can help in delivering accurate results depending on requirements and also develop the sustainability within project management performance. This study can bring the opportunity for incorporating more information related to practical and theoretical information due to which quality and clarity can be maintained successfully. Through this future scope, specific measurements can be identified properly for project managers that they can perform decision making processes effectively. It can bring stability and also develop the way of arranging sequence in project management performance that can allow in developing study skills. Development in big data knowledge can also help in understanding the processes of big data models and theories with respect to project management.

Bibliography

[Babiceanu and Seker, 2016] Babiceanu, R. F. and Seker, R. (2016). Big data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. *Computers in Industry*, 81:128–137.

[Badiezadeh et al., 2018] Badiezadeh, T., Saen, R. F., and Samavati, T. (2018). Assessing sustainability of supply chains by double frontier network dea: A big data approach. *Computers & Operations Research*, 98:284–290.

[Bailey, 2018] Bailey, G. (2018). The 5 vs for big data in project management. Available at: https://www.prosymmetry.com/resources/the-5-vs-for-big-data-in-project-management/.

[Bishop and Kuula-Luumi, 2017] Bishop, L. and Kuula-Luumi, A. (2017). Revisiting qualitative data reuse: A decade on. *Sage Open*, 7(1):2158244016685136.

[Bresnen, 2016] Bresnen, M. (2016). Institutional development, divergence and change in the discipline of project management. *International journal of project management*, 34(2):328–338.

[Caron et al., 2016] Caron, F., Ruggeri, F., and Pierini, B. (2016). A bayesian approach to improving estimate to complete. *International Journal of Project Management*, 34(8):1687–1702.

[Chae, 2015] Chae, B. K. (2015). Insights from hashtag# supplychain and twitter analytics: Considering twitter and twitter data for supply chain practice and research. *International Journal of Production Economics*, 165:247–259.

[Chou et al., 2016] Chou, J.-S., Hsu, S.-C., Lin, C.-W., and Chang, Y.-C. (2016). Classifying influential information to discover rule sets for project disputes and possible resolutions. *International Journal of Project Management*, 34(8):1706–1716.

[Doolun et al., 2018] Doolun, I. S., Ponnambalam, S., Subramanian, N., and Kanagaraj, G. (2018). Data driven hybrid evolutionary analytical approach for multi objective location allocation decisions: Automotive green supply chain empirical evidence. *Computers & Operations Research*, 98:265–283.

[Franková et al., 2016] Franková, P., Drahošová, M., and Balco, P. (2016). Agile project management approach and its use in big data management. *Procedia Computer Science*, 83:576–583.

[Henderson et al., 2016] Henderson, L. S., Stackman, R. W., and Lindekilde, R. (2016). The centrality of communication norm alignment, role clarity, and trust in global project teams. *International Journal of Project Management*, 34(8):1717–1730.

[Henrion, 2019] Henrion, M. (2019). Why most big data analytics projects fail. Available at: `https://pubsonline.informs.org/do/10.1287/orms.2019.06.08/full/#:~:text=In%202015%2C%20Gartner%20Research%20estimated,organizations%20was%20closer%20to%2085%25`.

[Hodgson and Paton, 2016] Hodgson, D. E. and Paton, S. (2016). Understanding the professional project manager: Cosmopolitans, locals and identity work. *International Journal of Project Management*, 34(2):352–364.

[Hong et al., 2018] Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., et al. (2018). Mixed methods appraisal tool (mmat), version 2018. *Registration of copyright*, 1148552:10.

[Iyer and Banerjee, 2016] Iyer, K. C. and Banerjee, P. S. (2016). Measuring and benchmarking managerial efficiency of project execution schedule performance. *International Journal of Project Management*, 34(2):219–236.

[Kartal et al., 2016] Kartal, H., Oztekin, A., Gunasekaran, A., and Cebi, F. (2016). An integrated decision analytic framework of machine learning with multi-criteria decision making for multi-attribute inventory classification. *Computers & Industrial Engineering*, 101:599–613.

[Kaur and Singh, 2018] Kaur, H. and Singh, S. P. (2018). Heuristic modeling for sustainable procurement and logistics in a supply chain using big data. *Computers & Operations Research*, 98:301–321.

[Larson, 2018] Larson, D. (2018). Agile project management and data analytics. *Data Analytics in Project Management*, 4(2):576–583.

[Legislation, 2018] Legislation (2018). Data protection act 2018. Available at: http://www.legislation.gov.uk/ukpga/2018/12/contents/enacted.

[Legislation, 2020] Legislation (2020). Understanding uk legislations. Available at: https://www.legislation.gov.uk/.

[Liu et al., 2016] Liu, S., Xia, F., Zhang, J., Pan, W., and Zhang, Y. (2016). Exploring the trends, characteristic antecedents, and performance consequences of crowdsourcing project risks. *International Journal of Project Management*, 34(8):1625–1637.

[Niazi et al., 2016] Niazi, M., Mahmood, S., Alshayeb, M., Qureshi, A. M., Faisal, K., and Cerpa, N. (2016). Toward successful project management in global software development. *International Journal of Project Management*, 34(8):1553–1567.

[Paquin et al., 2016] Paquin, J.-P., Gauthier, C., and Morin, P.-P. (2016). The downside risk of project portfolios: The impact of capital investment projects and the value of project efficiency and project risk management programmes. *International journal of project management*, 34(8):1460–1470.

[Sakka et al., 2016] Sakka, O., Barki, H., and Côté, L. (2016). Relationship between the interactive use of control systems and the project performance: The moderating effect of uncertainty and equivocality. *International Journal of Project Management*, 34(3):508–522.

[Saunders et al., 2015] Saunders, M. N., Lewis, P., Thornhill, A., and Bristow, A. (2015). Understanding research philosophy and approaches to theory development. *Research methods for business students*, 3(2):136–162.

[Silverman, 2020] Silverman, D. (2020). *Qualitative research*. Sage Publications Limited.

[Song et al., 2017] Song, M., Cen, L., Zheng, Z., Fisher, R., Liang, X., Wang, Y., and Huisingh, D. (2017). How would big data support societal development and environmental sustainability? insights and practices. *Journal of Cleaner Production*, 142:489–500.

[Tan et al., 2015] Tan, K. H., Zhan, Y., Ji, G., Ye, F., and Chang, C. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. *International Journal of Production Economics*, 165:223–233.

[van der Hoorn, 2016] van der Hoorn, B. (2016). Discussing project status with the project-space model: An action research study. *International Journal of Project Management*, 34(8):1638–1657.

[Wang et al., 2016] Wang, G., Gunasekaran, A., Ngai, E. W., and Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, 176:98–110.

[Wu et al., 2017] Wu, K.-J., Liao, C.-J., Tseng, M.-L., Lim, M. K., Hu, J., and Tan, K. (2017). Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties. *Journal of Cleaner Production*, 142:663–676.

[Zerbino et al., 2018] Zerbino, P., Aloini, D., Dulmin, R., and Mininno, V. (2018). Big data-enabled customer relationship management: A holistic approach. *Information Processing & Management*, 54(5):818–846.

[Zhang et al., 2016] Zhang, X., Bao, H., Wang, H., and Skitmore, M. (2016). A model for determining the optimal project life span and concession period of bot projects. *International Journal of Project Management*, 34(3):523–532.

[Zhang et al., 2017] Zhang, Y., Ren, S., Liu, Y., and Si, S. (2017). A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. *Journal of cleaner production*, 142:626–641.

[Zhong et al., 2016] Zhong, R. Y., Newman, S. T., Huang, G. Q., and Lan, S. (2016). Big data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives. *Computers & Industrial Engineering*, 101:572–591.