ECOSYSTEM DEVELOPMENT FRAMEWORK
FOR ENTERPRISE DATA SCIENCE & RESEARCH PROJECTS

by

Shitalkumar Rajendra Sukhdeve, DBA

DISSERTATION

Presented to the Swiss School of Business and Management Geneva

In Partial Fulfillment

Of the Requirements

For the Degree

DOCTOR OF BUSINESS ADMINISTRATION

SWISS SCHOOL OF BUSINESS AND MANAGEMENT GENEVA

June 2023

ECOSYSTEM DEVELOPMENT FRAMEWORK

FOR ENTERPRISE DATA SCIENCE & RESEARCH PROJECTS

by

Shitalkumar Rajendra Sukhdeve

APPROVED BY

<Chair’s Name, Degree>, Chair

Anna Provodnikova

<Member’s Name, Degree>, Committee Member

<Member’s Name, Degree>, Committee Member
Dr. Monika Singh, PhD.

RECEIVED/APPROVED BY:

<Associate Dean’s Name, Degree>, Associate Dean

DEDICATION

To my unwavering support system,

This thesis is dedicated to the incredible individuals, family, and friends who have stood by my side, offering their unwavering support and encouragement. Your belief in me has been my driving force. Thank you for being there throughout this journey.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to all those who have contributed to the successful completion of this thesis. Your support, guidance, and participation have been instrumental in shaping this research endeavor.

First and foremost, I extend my deepest appreciation to my mentor Monika Singh, whose expertise, patience, and invaluable insights have guided me throughout this journey. Your unwavering commitment to academic excellence and passion for the subject matter has been inspiring.

I would also like to acknowledge the participants who generously dedicated their time and shared their experiences for this research. Your willingness to participate in this study has contributed immensely to its depth and richness. Your valuable contributions have provided valuable perspectives and enriched the findings of this thesis.

Furthermore, I am grateful to the faculty members and researchers who provided valuable feedback and constructive criticism during the development of this thesis. Your expertise and dedication to scholarly inquiry have been instrumental in shaping and refining this work.

I appreciate my friends and family's unwavering support, understanding, and encouragement. Your belief in me, countless conversations, and words of motivation have kept me going, even in the face of challenges.

Finally, I want to express my gratitude to the academic community and all those whose works have laid the foundation for this research. Your ground breaking discoveries and scholarly contributions have paved the way for new insights and advancements.

To each and every person who has played a role, big or small, in this thesis, I extend my deepest thanks. Your support and contributions have made this research possible, and I am truly grateful for your presence in my academic journey.

With sincere appreciation,

Shitalkumar R. Sukhdeve

ABSTRACT

ECOSYSTEM DEVELOPMENT FRAMEWORK
FOR ENTERPRISE DATA SCIENCE & RESEARCH PROJECTS

Shitalkumar Rajendra Sukhdeve
2023

Dissertation Chair: <Chair's Name>
Co-Chair: <If applicable. Co-Chair's Name>

This research examines the development of an enterprise ecosystem for data science and research projects to unleash the full potential of data and foster innovation. The study proposes a comprehensive framework encompassing objective setting, firm performance evaluation, capability development, and strategic alignment. Industry experts and data science leaders from telecom, e-commerce, and financial sectors form the sample for interviews, which serve as the primary data collection method. Grounded theory is utilized for data analysis and theory development. The findings underscore the significance of objective setting, performance evaluation, capability development, and strategic alignment in optimizing processes and achieving superior outcomes. Organizations can cultivate innovation and gain a competitive edge by adopting this framework. The research contributes to effective data-driven decision-making and offers valuable insights for businesses operating in today's dynamic landscape.

Keywords: enterprise ecosystem, data science, research projects, objective setting, performance evaluation, capability development, strategic alignment, innovation, grounded theory, industry experts.

1 Table of Contents

List of Figures

CHAPTER I:

INTRODUCTION

1.1 Background

Enterprises have lots of data in their data warehouse. The legacy use of the data warehouse was to store data and generate some reports for the management. The global data warehousing market size was valued at $21.18 billion in 2019 and is projected to reach $51.18 billion by 2028, growing at a CAGR of 10.7% from 2020 to 2028 (Borasi, 2021). Enterprises poured money into data management software, which touched $73 billion in 2020 (Businesswire, 2021) . But still, the cost of data warehousing was not justifiable because data was not giving revenue by itself (Businesswire, 2021).

Enterprises started hiring Chief data scientists, data scientists, and machine learning experts and investing in big data technologies to meet their revenue targets or find new business opportunities (Sukhdeve, 2020).

But are things working as per expectations? Gartner Research predicted that 50 percent of chief data officers would fail (Gartner, 2019). Only 13 percent of data science and analytics projects reached production, meaning 7 percent ended without implementation (Roberts, 2017).

Many CDOs stated that they have a pilot of AI/ML, but it could not make it to production (Davenport, 2020). Harward's business review has quoted survey results that only 28 percent of participants (CDOs) accepted their role as successful and established (Davenport, 2020).

The NewVantage survey 2020 reported that 73.4 percent of firms accepted the business adoption of big data as a big challenge (NewVantage, 2020). Only 10 percent of startup companies with machine learning as their core business have reported revenue (Mckinsey, 2017).

A survey by Kaggle with 16000 data professionals reveals that 49.4 percent of participants registered challenges with dirty data. 41.6 percent of professionals reported a lack of data science talent in the organization. 37.2 percent of professionals said that lack of management and investment in data science teams are among the biggest challenges in organizations. 24.3 percent said that decision-makers do not use the outcomes of data science. Similarly,13.6 percent reported that data analysis findings are not integrated into decisions.

The gap in the use of data science outcomes and its use by decision-makers indicates the problem of the capability to use these insights and data trustworthiness. 30.2 percent of professionals consider data inaccessibility as one of the main problems of data science in enterprises, indicating the need for data governance (Kaggle, 2017).

Another survey from Kaggle shows that only 17.2 percent of candidates accepted that the organization has well-established Machine learning practices (Kaggle, 2020).

1.2 Research Problem

All this is because of 1) a Lack of understanding of data science. 2) Putting technology first rather than business problems and goals. 3) Lack of guiding principles or strategy to put all the necessary elements in place to execute the transformation needed for becoming a data science and research-driven organization.

Concisely, The research problem addressed in this study is the lack of a strategic ecosystem for developing, implementing, and utilizing data science and research projects in enterprises. The objective is to bridge this gap by proposing a comprehensive framework for ecosystem development. This framework will enable organizations in the telecom, eCommerce, and finance sectors to effectively plan, implement, develop, maintain, and leverage data science and research project outcomes.

1.3 Research Purpose and Questions

The primary aim of this research is to address the challenges enterprises encounter when it comes to developing and implementing data science and research projects within their ecosystems. The study seeks to provide valuable insights and recommendations that can facilitate the establishment of a strategic ecosystem capable of effectively utilizing the outcomes of data science and research projects in enterprises.

One of the main goals of this research is to bridge the existing gap in ecosystem development, specifically in data science and research projects in enterprises. To achieve this, the research has identified several objectives. Firstly, it aims to comprehensively understand the challenges and opportunities associated with ecosystem development for data science and research projects. This exploration will provide a nuanced perspective on the factors that influence the success of ecosystem development in enterprises.

Secondly, the research aims to identify the key factors that play a crucial role in the successful development of ecosystems within enterprises. By pinpointing these factors, the study will shed light on the elements that are pivotal for establishing an effective ecosystem.

Thirdly, the research endeavors to develop a comprehensive framework for ecosystem development that takes into account various essential components such as business objectives, technology, processes, and capabilities. This framework aims to provide a structured approach that integrates these elements to support the seamless development and functioning of the ecosystem.

Lastly, the research provides practical recommendations for implementing the developed framework and overcoming any barriers or challenges enterprises may face during ecosystem development. This research aims to facilitate the smooth and successful establishment of ecosystems in enterprises by offering actionable insights.

Several research questions have been formulated to guide the investigation toward achieving these objectives. These questions revolve around understanding the primary challenges faced by enterprises during the development of ecosystems for data science and research projects, identifying the critical success factors specific to ecosystem development in enterprises, devising strategies for designing and implementing a comprehensive framework for ecosystem development, and exploring the barriers to ecosystem development and effective methods to overcome them.

By thoroughly examining and addressing these research questions, this study provides a comprehensive exploration of the challenges, opportunities, and strategies involved in ecosystem development for data science and enterprise research projects. The insights and recommendations derived from this research are instrumental in guiding enterprises toward successfully establishing strategic ecosystems that leverage the full potential of data science and research outcomes.

1.4 Scope and Limitations

This research will primarily focus on enterprises operating in the telecom, eCommerce, and finance sectors, known for generating and managing large volumes of data. The findings and recommendations may also apply to other industries, but the selected sectors' specific contexts and challenges will be emphasized. It is important to note that the scope of this research does not include the technical aspects of data science algorithms or machine learning models but rather the strategic and organizational considerations in developing an ecosystem for data science and research projects.

1.5 Methodology

The research design for this study has employed a grounded theory approach, a qualitative research methodology focused on developing theories based on systematically gathered and analyzed data (Strauss & Corbin, 1994). This approach is well-suited for exploring complex phenomena and generating insights grounded in real-world data and experiences.

Semi-structured interviews were conducted with individuals who have expertise in implementing data science and research projects in enterprises. These interviews have provided rich insights into the challenges, opportunities, and strategies related to ecosystem development. Data coding was performed using an open, axial, and selective coding process to identify concepts, themes, and patterns. Memoing was used to record thoughts and interpretations during the analysis.

The data collection, analysis, and theoretical development processes were conducted simultaneously. Methodological consistency was ensured, and triangulation techniques will be employed to enhance the credibility and reliability of the research outcomes. The sample comprises individuals with at least three years of experience in relevant roles and industries.

1.6 Significance of the Study

This study holds significant importance for enterprises operating in the fields of telecom, eCommerce, and finance. By addressing the lack of a strategic ecosystem for data science and research projects, the proposed framework aims to bridge the gap and enable organizations to effectively plan, implement, develop, maintain, and leverage the outcomes of such projects.

The framework developed in this study will provide practical insights and recommendations for successfully establishing a strategic ecosystem that supports the utilization of data science and research project outcomes. It will help enterprises overcome challenges, identify key success factors, and navigate barriers associated with ecosystem development.

The findings of this research will contribute to the existing knowledge on ecosystem development for data science and research projects in enterprises. By exploring the challenges, opportunities, and strategies related to ecosystem development, this study will advance understanding in this domain and provide a basis for future research and development efforts.

1.7 Organization of the Thesis

This thesis is structured into six chapters, each serving a specific purpose to understand the research topic comprehensively. The following outlines the content of each chapter:

Chapter 1: The introduction, presents the background and context of the research and identifies the research problem, objectives, and research questions. It also highlights the significance of the study and provides an overview of the thesis structure.

Chapter 2: Literature Review In the literature review chapter, relevant theories, models, frameworks, and studies related to ecosystem development for data science and enterprise research projects are critically reviewed. It identifies existing gaps, synthesizes existing knowledge, and provides a theoretical foundation for the research.

Chapter 3: Methodology The methodology chapter describes the research design, data collection procedures, sample selection, and data analysis techniques employed in the study. It explains the grounded theory approach as the chosen methodology and justifies its suitability. The chapter also addresses any ethical considerations.

Chapter 4: Results In the results chapter, the findings obtained from the data analysis are presented. It includes a detailed analysis of the collected data and identifying key concepts, themes, and patterns related to ecosystem development. The results are presented clearly and organized, supported by relevant quotes and examples from the data.

Chapter 5: Discussion, The discussion chapter interprets and analyzes the results in light of the research objectives and existing literature. It provides an in-depth analysis of the findings, explores their implications, and relates them to the research questions. The chapter also discusses any limitations of the study and suggests areas for further research.

Chapter 6: Summary, Implications, and Recommendations The final chapter summarizes the study's main findings and restates the research objectives. It discusses the implications of the research in the context of ecosystem development for data science and research projects in enterprises. Additionally, this chapter offers practical recommendations based on the research outcomes and suggests potential avenues for future research.

1.8 Summary

In this chapter, the research problem of the lack of a strategic ecosystem for the development, implementation, and utilization of data science and research projects in enterprises was addressed. The objective of this study is to propose a comprehensive framework for ecosystem development that bridges this gap and enables organizations to effectively leverage data science and research project outcomes.

The research purpose and questions were outlined, focusing on understanding the challenges and opportunities associated with ecosystem development, identifying key success factors, developing a comprehensive framework, and overcoming barriers to ecosystem development in enterprises. The research design utilizing a grounded theory approach was described, emphasizing iterative steps such as data collection, coding, memoing, and analysis.

The population and sample for the study were discussed, with an initial focus on a homogenous sample progressing towards a heterogenous sample as the theory develops.

Purposeful and convenience sampling are employed to select individuals with expertise in implementing enterprises' data science and research projects.

Data collection procedures involve conducting semi-structured interviews with the selected participants. The sample size was determined based on data saturation, ensuring enough participants are included to reach a point where new information no longer significantly contributes to the emerging theory. Triangulation techniques enhance the credibility and reliability of the research outcomes.

Data analysis will follow a rigorous and iterative process guided by grounded theory principles. The three coding phases (open, axial, and selective) will be employed, with memoing facilitating the recording of reflections and ideas. Theoretical sampling will guide the exploration of emerging concepts and themes, ensuring constant refinement and expansion of the theoretical framework.

The thesis structure was outlined, encompassing Introduction, Literature Review, Methodology, Results, Discussion, and Summary chapters. Each chapter serves a specific purpose in providing a comprehensive understanding of ecosystem development for data science and research projects in enterprises.

In summary, this research aims to contribute to the field of enterprise ecosystem development by proposing a comprehensive framework and providing insights and recommendations for successful implementation. The subsequent chapters will delve into the literature, methodology, results, discussion, and implications, further advancing our understanding in this area.

CHAPTER II:

REVIEW OF LITERATURE

2.1 Introduction

This chapter comprehensively reviews the existing literature on enterprise development for data science and research projects. The review aims to explore the current state of research, identify key findings from previous studies, and highlight gaps in the literature that warrant further investigation. By examining relevant studies, frameworks, and methodologies, this literature review sets the foundation for the subsequent chapters of this thesis.

2.2 Enablers of Data Science Adoption

Various enablers influence the adoption of data science in organizations. (Behl, et al., 2019) conducted a study on the adoption of big data analytics by startups and identified enablers such as technical support from vendors, management attitude, IT infrastructure, and training and development of employees. Similarly, (Rene Abraham, 2019) found that factors like organizational culture, leadership support, data quality, and data governance play crucial roles in facilitating successful data science adoption. However, further research is needed to explore the specific mechanisms through which these enablers impact adoption (Behl, et al., 2019) (Rene Abraham, 2019).

2.3 Data Science Strategy Adoption

Adopting data science strategies in organizations is a complex process that requires careful planning and implementation. (Rajesh, et al., 2022) Conducted a systematic literature review on data science strategy adoption and proposed a conceptual framework comprising three layers: input, enterprise data science constituents, and outcome. This framework provides a holistic view of the factors influencing strategy adoption, including resources, barriers, enablers, and strategic business values. However, the framework requires validation in real-world settings and with expert input. Additionally, further research is needed to explore specific components of data science strategies, such as data visualization strategies, machine learning, and deep learning strategies (Rajesh, et al., 2022) .

2.4 Big Data-Business Strategy Research

The intersection of big data and business strategy has received significant attention in research. (Ciampi, et al., 2020) conducted a bibliometric analysis and systematic literature review, mapping the big data-business strategy research field. They identified four thematic clusters: big data and supply chain strategy, personalization and co-creation strategies, strategic planning and value creation pathways, and the relationship between big data and knowledge management. However, further research is required to address potential biases in the selection process and explore these clusters' applications in various organizational contexts (Ciampi, et al., 2020).

2.5 Value Creation in Big Data Scenarios

Value creation is a critical aspect of leveraging big data in organizations. (Elia, 2019) proposed a framework for evaluating value dimensions in Big Data scenarios. Their framework identified four dimensions: Informational Value, Transactional Value, Transformational Value, and Strategic Value. These dimensions are associated with various organizational benefits, such as cost reduction, operational efficiency, competitive advantage, and innovation. However, the framework lacks granularity regarding actionable steps for developing the ecosystem for data science and big data. Future research could focus on identifying specific actions and strategies that organizations can implement to leverage these value dimensions effectively (Elia, et al., 2020).

2.6 Evaluating Analytic Maturity

Assessing the maturity of analytics capabilities in organizations is crucial for effective data science development. (Grossman, 2018) developed the Analytic Processes Maturity Model (APMM) for this purpose. The APMM identifies five levels of capability, ranging from basic reporting to strategy-driven analytics. This model provides organizations with a framework to assess their current analytics maturity and identify areas for improvement. However, it falls short of providing specific guidelines for setting up an enterprise data science ecosystem. Future research should focus on developing comprehensive guidelines that consider factors beyond analytic maturities, such as data governance, talent management, and technology infrastructure (Grossman, 2018).

2.7 Analytics Adoption in Organizations

(LaValle, et al., 2011) conducted a study on analytics adoption in organizations and proposed a framework with three stages: aspiration, experience, and transformation. They examined various capabilities within each stage, such as motive, functional proficiency, business challenges, and data management. The study also provided recommendations for implementing analytics projects, including starting with the most significant opportunities, embedding insights into actions, and adding new capabilities while maintaining existing ones. However, further research is needed to validate the use of these recommendations across different organizational contexts and to assess the long-term impact of analytics adoption (LaValle, et al., 2011).

2.8 Enterprise Data Science Ecosystem

A Comprehensive Framework (Reddy, 2022) conducted a study proposing a comprehensive framework for building an enterprise data science ecosystem. Their framework encompasses critical components such as data governance, infrastructure, talent management, collaboration mechanisms, and organizational culture. By considering these factors holistically, organizations can establish a robust ecosystem that fosters data science and research projects. Further research can focus on validating and refining this framework in different organizational contexts (Reddy, 2022).

2.9 Agile Approaches for Data Science Projects

Agile methodologies have gained attention in data science project management. (Dastgerdi & Gandomani, 2021) investigated the application of Agile principles and practices in data science projects and identified key challenges and success factors. They highlighted the importance of iterative development, cross-functional teams, stakeholder collaboration, and continuous feedback in achieving successful outcomes. However, more research is needed to explore specific Agile frameworks and methodologies tailored to the unique characteristics of data science projects (Dastgerdi & Gandomani, 2021).

2.10 Ethical Considerations in Enterprise Data Science

Ethical considerations play a crucial role in enterprise data science. (Georgieva, 2022) examined the ethical challenges and implications of data science in organizational settings. They emphasized the need for ethical guidelines, transparency, and accountability throughout the data science lifecycle. Future research should develop frameworks and guidelines that promote responsible and ethical data science practices in organizations (Georgieva, 2022).

2.11 Integration of Artificial Intelligence in Enterprise Data Science

Integrating artificial intelligence (AI) technologies in enterprise data science has transformative potential. (Rashmi Yogesh Pai, 2022) conducted a study on the adoption of AI in data science projects and identified critical success factors, including data quality, algorithm selection, and interpretability of AI models. They also highlighted challenges related to bias, privacy, and explainability. Further research can explore AI integration frameworks, address ethical concerns, and investigate the impact of AI on organizational performance (Rashmi Yogesh Pai, 2022).

2.12 Industry-specific Applications of Enterprise Data Science

Enterprise data science has industry-specific applications that require tailored approaches. (Abdelbaki, et al., 2023) investigated the adoption of data science in the healthcare sector and proposed a framework considering domain knowledge, data integration, analytics models, and decision-making processes. Similar industry-specific studies can be conducted in other sectors, such as finance, manufacturing, and retail, to develop context-specific enterprise data science frameworks (Chen et al., 2022).

2.13 Key Findings and Gaps in the Literature

The reviewed studies reveal several key findings regarding enterprise development for data science and research projects. These findings emphasize the importance of comprehensive frameworks, such as (Reddy, 2022) ecosystem framework, that consider various components such as data governance, infrastructure, talent management, and collaboration mechanisms.

However, despite the valuable insights existing research provides, there are still notable gaps in the literature. For instance, there is a need for further exploration of Agile methodologies and frameworks explicitly tailored for data science projects (Dastgerdi & Gandomani, 2021) . More research is required to address ethical considerations in enterprise data science and develop guidelines to promote responsible practices (Georgieva, 2022). Furthermore, the integration of AI technologies and industry-specific applications of enterprise data science requires more in-depth investigation to fully understand their implications and benefits (Rashmi Yogesh Pai, 2022).

This chapter has provided an overview of the existing research on enterprise development for data science and research projects. The reviewed studies have shed light on key components of an enterprise data science ecosystem, the application of Agile methodologies, ethical considerations, the integration of AI, and industry-specific applications. While these studies have contributed valuable insights, gaps still require further exploration and investigation.

By considering the latest research and addressing these gaps, this thesis aims to develop a comprehensive framework for enterprise development in the context of data science and research projects. The subsequent chapters will delve deeper into the methodology, findings, and recommendations, ultimately contributing to advancing knowledge in this field and providing practical guidelines for organizations seeking to establish thriving enterprise data science ecosystems.

CHAPTER III:

METHODOLOGY

3.1 Overview of the Research Problem

Enterprises today are investing heavily in data science and artificial intelligence, with the global data warehousing market projected to reach $51.18 billion by 2028 (Borasi, 2021). However, despite these investments, many organizations struggle to achieve significant revenue and successful outcomes from their data initiatives (Businesswire, 2021; McKinsey, 2017). Challenges such as a lack of understanding of data science, prioritizing technology over business goals, and the absence of a guiding ecosystem strategy hinder the effective utilization of data science and research projects in enterprises (Gartner, 2019; Roberts, 2017; NewVantage, 2020; Sukhdeve,2020).

The research problem addressed in this study is the lack of a strategic ecosystem for developing, implementing, and utilizing data science and research projects in enterprises. The objective is to bridge this gap by proposing a comprehensive framework for ecosystem development. This framework will enable telecom, eCommerce, and finance organizations to effectively plan, implement, develop, maintain, and leverage data science and research project outcomes.

3.2 Research Purpose and Questions

This research aims to address the challenges enterprises face in developing and implementing data science and research projects within their ecosystems. The study aims to provide insights and recommendations for successfully establishing a strategic ecosystem that supports the utilization of data science and research project outcomes in enterprises.

This research aims to fill the ecosystem development gap for data science and research projects in enterprises. Specifically, the objectives are:

1. To understand the challenges and opportunities associated with ecosystem development for data science and research projects.
2. To identify the key factors influencing the success of ecosystem development in enterprises.
3. To develop a comprehensive framework for ecosystem development that integrates business objectives, technology, processes, and capabilities.
4. To provide recommendations for implementing the framework and overcoming barriers to ecosystem development in enterprises.

To achieve these objectives, the following research questions will guide the investigation:

1. What are the main challenges faced by enterprises in developing an ecosystem for data science and research projects?
2. What are the key success factors for ecosystem development in enterprises?
3. How can a comprehensive framework for ecosystem development be designed and implemented?
4. What are the barriers to ecosystem development in enterprises, and how can they be overcome?

These research questions will comprehensively explore the challenges, opportunities, and strategies related to ecosystem development for enterprises' data science and research projects.

3.3 Research Design

The research design for this study has employed a grounded theory approach. Grounded theory is a qualitative research methodology that aims to develop theories based on systematically gathered and analyzed data (Strauss & Corbin, 1994). It is particularly suitable for exploring complex phenomena and generating new insights and theories grounded in real-world data and experiences.

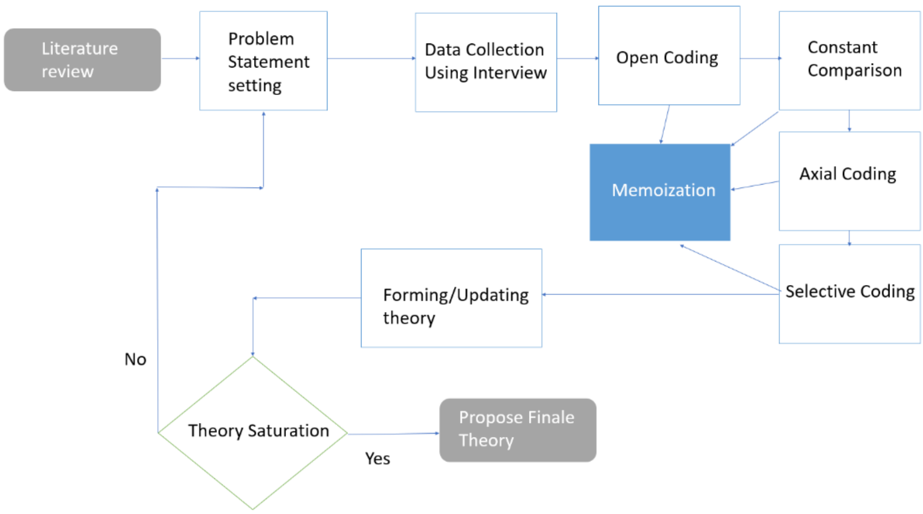

Figure 1. Grounded Theory flowchart

Figure 1 depicts the research design involving iterative and interrelated steps, including data collection, coding, memoing, constant comparison, and writing (Dick, 2005). Semi-structured interviews are conducted with individuals with expertise in implementing enterprises' data science and research projects. The interviews provide rich insights into the challenges, opportunities, and strategies related to ecosystem development. Notetaking during the interviews captures important details and emerging ideas.

Data coding is conducted using an open, axial, and selective coding process (Strauss & Corbin, 2015). The data is systematically analyzed to identify concepts, themes, and patterns related to ecosystem development. Memoing records thoughts, interpretations, and ideas that arise during the analysis process. The iterative nature of the grounded theory approach allows for constant refinement and modification of the research questions and theoretical framework based on emerging findings.

The data collection, analysis, and theoretical development processes are conducted simultaneously. The data is synthesized and organized to develop a coherent and comprehensive theoretical framework for ecosystem development. The findings are logical and structured, supported by relevant quotes and examples from the data.

Throughout the research, attention has been given to methodological consistency, ensuring the integrity and validity of the findings. Triangulation techniques, such as using multiple data collection through interviews, cross-validation of findings, and bias removal using throughout the cross-validation process, have been employed to enhance the credibility and reliability of the research outcomes.

In summary, the grounded theory approach provides a robust and systematic methodology for exploring the complexities of ecosystem development for data science and research projects in enterprises. It ensures the practical applicability and relevance of the proposed framework by grounding it in real-world data and experiences.

3.4 Population and Sample

This study focuses on the process level rather than individual participants. The research begins with a homogenous sample of individuals and progresses toward a heterogenous sample as the theory develops (Strauss & Corbin, 2015). Purposeful sampling, which strategically selects information-rich cases, will be used in conjunction with convenience sampling (Patton, 2002).

A group of individuals identified who possess expertise in implementing data science and research projects in enterprises. The sample consists of individuals with at least three years of experience as data scientists, researchers, or project managers in handling data science projects. They must have implemented or managed at least one-use case utilized by the business and should be employed in industries that deal with substantial amounts of data, such as telecom, eCommerce, and finance.

Initial participants are selected through convenience sampling, utilizing personal contacts such as LinkedIn connections and peers' recommendations. Additional interview participants are identified through purposeful sampling by requesting recommendations from individuals in the initial sample. The sample size has been determined based on data saturation, which occurs when no new properties, dimensions, or relationships emerge during analysis (Strauss & Corbin, 1998).

In addition to purposeful sampling, theoretical sampling has been employed to find manifestations of the theoretical constructs of interest and further enrich the categories and themes (Patton, 2002).

3.5 Data Collection Procedures

Data collection for this research has involved conducting semi-structured interviews with individuals who fulfill the abovementioned criteria.

Candidates have been contacted through emails and inbox messages to explain the purpose of the research, invite their participation, and schedule interview sessions. With purposeful sampling, requested recommendations from initial participants to identify additional interviewees who can provide valuable insights and fill any potential gaps in the data.

The sample size has been determined based on the principle of data saturation. Saturation occurs when no new properties, dimensions, or relationships emerge during the analysis process (Strauss & Corbin, 1998). As the grounded theory approach emphasizes theoretical development based on data, there are no predetermined guidelines or tests of adequacy for estimating the sample size required to reach saturation (Fast, 2021). The aim is to include enough participants to reach a point where new information no longer significantly contributes to the emerging theory.

To complement purposeful sampling, theoretical sampling has been employed. Theoretical sampling involves selecting cases that manifest the theoretical construct of interest and provide opportunities for elaboration and examination of the construct and its variations (Patton, 2002). By comparing similarities and differences among a broad range of cases, the research aims to add variation and density to the emerging categories and guide further sampling from information-rich sources (Strauss & Corbin, 2015).

3.6 Data Analysis

Data analysis in this research involves a rigorous and iterative process guided by the principles of grounded theory. The analysis encompasses three phases of coding: open coding, axial coding, and selective coding (Corbin & Strauss, 2015). Open coding involves the initial exploration of the data, identifying concepts, and assigning descriptive labels to data segments. Axial coding focuses on establishing relationships between categories and subcategories, while selective coding aims to integrate and refine emerging theory.

Throughout the analysis, memoing will be used to record reflections, interpretations, and ideas that emerge during the coding process. Memos provide a space for capturing and organizing thoughts, allowing for developing connections and theoretical insights. The constant comparison method, which involves comparing new data with previously analyzed data, has been employed to identify similarities and differences, refine categories, and develop theoretical explanations (Corbin & Strauss, 2015).

Theoretical sampling guides the data collection and analysis process, allowing for further exploration of emerging concepts and themes. This iterative approach ensures that the theoretical framework is refined and expanded as new insights and patterns emerge from the data.

Two additional steps have been taken to ensure the integrity and validity of the theory. First, the raw data is revisited to assess the explanatory power of the emerging theory across most cases (Strauss & Corbin, 2015). This step examines the fit between the theory and the data, ensuring that the more significant concepts or categories apply to each case. Second, follow-up interviews have been conducted with the participants to gather feedback on the developed framework. This feedback loop between the researcher and the interviewees further enhances the theory's validity and applicability to real-world contexts (Fast, 2021).

The research findings are presented comprehensively and organized, showcasing the key concepts, categories, and relationships that have emerged from the data analysis. Visual representations, such as diagrams and models, have been used to illustrate the theoretical framework and enhance its understanding.

3.7 Ethical Considerations

Ethical guidelines and principles have been followed throughout the research process. Informed consent has been obtained from all participants, ensuring confidentiality and privacy of their information. The data collected has been securely stored and anonymized to protect the participants' identities. The research adheres to ethical guidelines provided by the SSBM institution.

3.8 Limitations

This study has some limitations that should be considered. Firstly, the generalizability of the findings may be limited due to the focus on specific industries and contexts. Secondly, the sample size has been determined by data saturation, which results in a relatively small sample. However, the iterative nature of the grounded theory approach allows for in-depth exploration and understanding of the phenomena under investigation. Finally, the research depends on the participant's willingness to provide accurate and detailed information.

Despite these limitations, the research provides valuable insights and practical recommendations for enterprises aiming to develop and implement data science and research projects within their ecosystems.

3.9 Summary

This chapter provided an overview of the research methodology and design employed to develop a framework for ecosystem development for data science and enterprise research projects. The grounded theory approach is selected as the methodology due to its ability to generate theory inductively from actual data and its emphasis on practical applicability.

The chapter discussed the research purpose and questions, highlighting the aim of filling the gap in strategic ecosystem development for data science and research projects. The research design, including the grounded theory approach, purposeful sampling, and theoretical sampling, was explained in detail. Data collection procedures involving semi-structured interviews and sample size determination through data saturation were described.

The chapter also outlined the data analysis process, including the three coding phases, memoing, and constant comparison. The steps taken to ensure the integrity and validity of the emerging theory, such as revisiting the raw data and conducting follow-up interviews, were discussed.

By employing this rigorous research methodology, the study aims to develop a framework that addresses the challenges and gaps in ecosystem development for data science and research projects in enterprises. The subsequent chapters will delve into the data collection and analysis process, presenting the findings and insights from the study.

CHAPTER IV:

RESULTS

The study aims to understand how industry data leaders deliver data science and research projects. It explores the process and strategies data leaders follow to develop an ecosystem in the organization that makes them deliver these projects. It presents a comprehensive process of developing an ecosystem for data science and research project delivery. This work used rigorous grounded theory procedures, such as constant comparative analysis and data saturation, to develop a theory inductively from data collected through expert interviews.

This chapter discusses a comprehensive overview of the grounded theory research methodologies used by the researcher, along with the study's conclusions and an ecosystem development framework for data science and research initiatives. Based on how well the results show validity, credibility, and applicability, the reader should be able to determine whether the grounded theory technique has been successfully applied.

4.1 Initiating the Study

To understand how data leaders in the industry approach developing an ecosystem for data science and research, 21 individual data experts and leaders were interviewed. The participants came from telecom, e-commerce, and finance organizations and countries like the UK, US, Japan, India, Indonesia, Singapore, Canada, Dubai, and Malaysia.

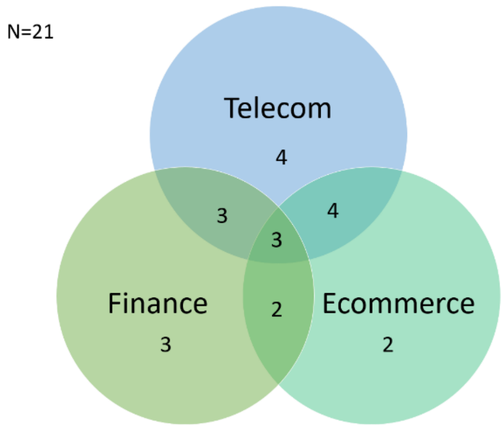

Figure 2. Population Distribution

Figure 2 shows that they also have experience working in multiple industries. They were selected through the researcher's network and personal experience as individuals who satisfied the study's criteria and had relevant experience developing and delivering data science and research projects.

The entire research procedure, including note-taking, coding, and memo-writing, was carried out throughout the study. Specifics on how the study approach was carried out are provided in the following section.

4.2 Executing the Research Method

The only qualitative approach that places the research process at its top (Strauss & Corbin, 2015) and has the explanatory power required to identify the cognitive and behavioral stages involved in creating a new framework for ecosystem development and to create a comprehensible theoretical justification is grounded theory. Grounded theory was therefore chosen as the research method for this study.

To inform the theory-building process, the grounded theory design sought out the expertise and viewpoint of professionals skilled in data science and the implementation of research projects in enterprises. Data were gathered using Zoom, a video conferencing tool, during semi-structured, in-depth interviews. Data were gathered and analyzed simultaneously. To create the categories and theoretical linkages particular to this research approach, the process involved interviewing, coding, memoing, sorting, and writing in a continuous comparison manner. New information and conditions emerged to challenge and reinforce the researcher's interpretation of the data.

Once the theoretical abstractions and the actual data fit well, a visual model representing the essential categories and concepts was created. A second phase of the study was used as follow-up interviews to get participant feedback on the model and ensure the theory was properly assimilated and validated.

4.3 The sample

As stated above, data was collected from 21 participants, mainly from telecom, e-commerce, and finance-like data-heavy organizations in India, Indonesia, Singapore, Belgium, Malaysia, Dubai, etc. The participants are experts in data science and research project development and implementation based on the criteria mentioned in Chapter III. The participants meet the criteria- 1) Individuals must have at least 3 years of experience working as a data scientist, researcher, or project manager handling data science projects. 2) They must have implemented or managed at least one-use case that is used by the business. 3) They should be in the industries which handle big data like telecom, eCommerce, etc.

Eight people were included in the original sample of participants who were identified through the researcher's personal connections. An extra 30 qualified interview candidates were found using a combination of theoretical sampling and purposeful sampling. The researcher requested suggestions from individuals who would fulfill the study parameters and contribute new perspectives to the researcher's understanding of developing concepts from current research participants and other industry contacts. Not everyone invited to participate in the theoretical and purposeful sample could devote the necessary time to the interview procedure and opt-out. Nonetheless, 13 of the 30 people who completed the requirements and used the theoretical sampling technique agreed and participated. 21 expert interviews were thus performed in all.

As presented in Table 1, participants had a range of educational and professional backgrounds, including those with MBAs, PhDs, Master in Technology from computer science degrees, telecommunication engineering degrees, finance degrees, and at least three years of experience. 18 men and 3 women made up the 21-person sample.

Candidates in the study held prominent positions in their organizations. Each candidate interviewed had launched at least one successful data science or research project that a business had adopted. Candidates held positions ranging from senior data scientist to chief data scientist. These people were in charge of projects from conception to deployment and commercialization.

Those interviewees who had experience in several different industries and had built an organization's data science and research ecosystem from the ground up provided the most detailed information. These cases allowed the researcher to investigate several conditions that led to the discovery in greater depth.

Table 1. Data Sample

No. of Participants	Experience Level	Education Level	Location
5	More than 10 years	PhD	United States, India, UK, Indonesia
2	5 to 10 years	Master's	United Kingdom, India
3	More than 10 years	Master's	Canada, Indonesia, India
2	More than 10 years	Master's	India
1	More than 10 years	Master's	Singapore
3	More than 10 years	Master's	Singapore
2	More than 10 years	Master's	United States
3	More than 10 years	Master's	India, UK, Indonesia

4.4 Data collection

The data collection process for the interviews consisted of two phases. In the first phase, interviews were conducted over three months, while in the second phase, they were conducted over five months, totaling eight months of data collection. The interviews were carried out virtually using video conferencing technology, specifically Zoom. Before collecting data, the researcher put in an extensive effort to develop the line of inquiry. This was followed by a pilot study involving two interviews, which helped the researcher receive feedback from a second reader employed in the study and a sample representative. The sequencing and critical nature of each interview question were also reviewed. Each question was then analyzed about the literature to ensure that the line of inquiry addressed all relevant areas of literature.

The pilot study was initiated and followed the grounded theory research design and strategy described in Chapter 3 to ensure its authenticity. The researcher reached out to two individuals who matched the sample criteria through email and invited them to participate in the study. The email template used for the initial contact is as follows:

Dear [Participant name],

I hope this message finds you well. I am writing to inquire if you would be interested in participating in my doctoral research by granting me an interview. My dissertation focuses on developing an "Ecosystem development framework for enterprise data science and research projects." With your experience in leading data science and research projects in [xyz company or developing projects], you would be an ideal participant for my study.

My research aims to investigate how data science leaders create ecosystems for data science and research projects in enterprises. I am seeking interview participants who meet at least two of the following criteria: 1) a minimum of 3 years of experience in data science or research areas, 2) the successful execution of at least one data science project, and 3) experience leading a team of data scientists or researchers.

I value your time and will ensure our interview is as efficient as possible. At your convenience, the research will require a few hours of your time over the next few months. The initial interview is expected to take between 50 and 120 minutes, followed by a shorter follow-up interview to obtain feedback on the model as it emerges.

I understand that confidentiality may be a concern. I assure you that I will not collect proprietary information during the interview, and all personal identifying information will be removed from the results.

Your participation in this study would be precious and contribute to a substantial area of the scholarly literature. Moreover, it would help establish an ecosystem development framework for data science and research projects that is relevant and valuable to industry leaders. I would be delighted to present my final results to you and/or your organization.

Please let me know if you are interested in participating. I am happy to discuss any questions or concerns you may have over a quick phone call before committing.

Thank you for considering my request.

Best regards, [Researcher's Name]

Each of the two individuals agreed to participate in the study and was requested to schedule a suitable time for the interview using the Google Calendar application. They were also asked to sign a consent form before the interview. The following email template was used for scheduling and obtaining consent from the participants:

Dear [Participant's Name],

Thank you very much for agreeing to participate in my dissertation research. Your willingness to assist me is greatly appreciated. To proceed with scheduling our interview, please let me know a convenient time for you by responding to this email. I have blocked off Saturdays on my calendar for conducting interviews, but if that day of the week is not suitable for you, kindly inform me.

Once we have established a suitable date and time, I will send you a Zoom link for the meeting, which you can add to your calendar. Before the interview, please take the time to read the Informed Consent Form that I have attached to this email. Please indicate your agreement to participate in the study by checking the box at the bottom of the form and typing your name.

Thank you again for your cooperation. I look forward to our conversation.

Best regards, [Researcher's Name]

To initiate each interview, the researcher utilized a consistent script to build a connection with the participant, maintain confidentiality, and ask for consent to record the session. Below is the introductory script used by the researcher:

"Thank you for taking the time to speak with me today. As mentioned, my doctoral research focuses on developing an ecosystem development framework for data science and enterprise research projects. Through my research and experience, I have discovered that the success rate of such projects is low, and their adoption by businesses is difficult. My research aims to create a framework or guidebook to help organizations integrate data

science and research into their business, from management-level strategic preparation to execution-level readiness.

As a data leader in your organization, I am interested in understanding the various initiatives and strategic considerations you believe are essential for developing such an ecosystem. I want to assure you that your responses will be kept confidential, and I will protect your identity by assigning a pseudonym and removing any personally identifiable information during data analysis. Do you have any questions about confidentiality?

Also, may I record our conversation today? This will allow me to focus on listening rather than taking notes.

Now, let's begin. I have a few open-ended questions to start with, and then we will move on to some directed questions based on insights from previous interviews."

The researcher utilized a transcription service through **https://beecut.com/speech-to-text-online** and Sembly to transcribe the recorded interviews. Atlasti and Chatgpt were used for note-taking, coding, and memoing. The pilot study progressed seamlessly without major alterations to the research protocols or line of questioning. Consequently, the results from the pilot study were integrated into the final analysis, and the researcher proceeded with scheduling and conducting additional interviews concurrently with the data analysis.

The data collection for the initial phase involved conducting one-on-one interviews that lasted between 40 to 60 minutes each. These interviews followed a semi-structured format and utilized the same data collection and analysis process outlined in Chapter 3 and performed during the pilot study.

The interviews were scheduled, conducted, coded, and reflected upon simultaneously using detailed notes and memos—the theory-building process employed three phases of coding: open, axial, and selective coding. During the open coding phase, the researcher examined the data sets and searched for patterns or relationships among the participants' experiences.

The emerging concepts were then divided into natural categories and subcategories, which changed as new data sets were added. In the axial coding phase, links were established between categories as the researcher examined the various conditions, action-interaction effects, and consequences in the data. Finally, selective coding was utilized to determine the core category and unify all other categories.

Throughout the process, the researcher utilized memos and diagrams to facilitate the understanding of the intricate relationship between variables present in the data. The participants exhibited a diverse lexicon and frequently used metaphors to articulate their experiences, which proved valuable for the researcher to comprehend the data from various perspectives.

To reduce bias and enhance precision, a conversational AI tool named Chatgpt was used at different stages of grounded theory analysis and cross-validation through participant feedback. Coding and axial coding were conducted using this tool, and relationships were generated by executing and regenerating responses multiple times. This aided the researcher in refining the data narrative and detecting any inconsistencies in the logic.

Data saturation was reached by Interview 9. The diversity of the sample, which included participants with experience in multiple industries, such as data scientists, managers, and researchers in enterprises from different locations, presented several dimensions to consider and verify a repeatable pattern. Once saturation was reached, 12 more interviews were conducted to check for any deviations from the other cases and gain additional depth of perspective.

Once the first data collection phase was completed, the researcher developed a new model of the ecosystem development framework for enterprise data science and research projects, including supporting proposition statements. The second round of interviews was initiated to collect feedback on the model. Participants were sent a follow-up email for a shorter interview lasting 30 minutes to 1 hour.

Two days before the call, interviewees received a visual of the model and supporting proposition statements for review. During the follow-up interview, the researcher gave a brief overview of the model, and respondents were asked to discuss which aspects of the model fit or did not fit their experience with ecosystem development. The researcher also asked additional probes to determine whether any items should be added, removed, or rearranged on the model to rank the variables and whether they knew how ecosystems not currently represented in the model could be developed. This feedback loop helped to refine the researcher's understanding of the relationship between concepts and to address existing biases.

During the second phase, feedback was requested from 21 interviewees, out of which 19 agreed to participate. One participant who could not schedule a Zoom meeting provided feedback via email.

4.5 Data analysis and findings

According to Corbin and Strauss (2015), the process of developing a theory using coding was followed. This involved three phases of data analysis: open, axial, and selective coding. In the open coding phase, the researcher identified lower-level concepts and discovered patterns and relationships between them by constantly comparing within and across data sets. A code was assigned to each line of every interview. For instance, the sentence "institutional barriers" was identified as a code for the following quote:

“data science doesn't work with middle management. It works only when the top management of not only your department but other stakeholder departments is already involved if you work with only middle management, many of those projects don't work.”

After coding each interview line by line, the researcher added each unique code to a master code list. Codes from subsequent interviews were also added and compared to previous interviews to identify similarities and differences. This iterative process continued until saturation was reached and all unique codes had been explored under various conditions. As a result of this process, natural categories or "buckets" for grouping the codes began to emerge. Grouping the codes in this way provided a foundation for the next phase of coding.

The researcher employed memoing techniques during the coding process to capture and process insights gained from each interview. To begin with, the memo would encapsulate the interviewee's viewpoint. Observational notes were made to record the critical incidents that led to their experiences. Lastly, theoretical notes were noted down to interpret and analyze the meaning of these experiences. These notes served to capture any surprising or conflicting statements made during the interviews, compare and contrast statements between interviews, and identify areas for further investigation. As the analysis advanced, summary memos were used to recapitulate previous analyses, while linking memos were created to document hypotheses regarding the relationship between multiple codes.

The subsequent phase of coding, known as axial coding, required a reassessment of the results obtained from open coding, intending to establish connections between emerging categories by critically evaluating the data under varying conditions. The master code list contained the initial categories, unique codes, and frequency of each code across the dataset. The researcher scrutinized dominant codes and emerging themes to understand the relationship between categories in diverse situations. With the help of memoing and diagramming techniques, a consistent pattern or sequence of events emerged from the data, allowing the researcher to establish links and re-label the initial categories to better reflect the action-interaction effects evident in the data. Here is one example of how three categories, BDA Management Capability and BDA Technology Management, BDA Talent Management, were linked with the sub-dimension of BDA Capability.

Tag: BDA Management Capability

Quote: “In terms of the data infrastructure and. I am working with teams of the twelve single data platforms and some investments currently happening. I would say the IT department manages the big data platform.”

Tag: BDA Technology Management

Quote:” We considered here as the users of the data more than a team that manages the data directly. So we have a separate team to manage the data and which is in the operational data store.”

Tag: BDA Talent Management

Quote: “Here in my bank, hiring data science. It is not given particular attention. Also, I just it's just like hiring any other employee. Is there is it it has become. You can see almost like a mainstream road, which was not the. The thing before, right where data scientists were rare to find that it was. Uh, what? As I said, extraordinary attention was is given to hiring data science.”

The instances above demonstrate how the BDA Management Capability, BDA Technology Management, and BDA Talent Management are interconnected. This theme was prevalent in the majority of interviews and also concerned other sub-dimensions of Capability Development.

In the final stage of coding, selective coding was employed to scrutinize the causal connections among categories, identify the significant category of the phenomenon, and integrate the categories around it. This phase involved using diagramming extensively to unify categories and identify the primary theme of the study.

When participants were asked to name the ecosystem development influencing factors, they mapped to the four-dimensional theme: business objective setting, firm performance evaluation, capability development, and strategic alignment (Figure 3).

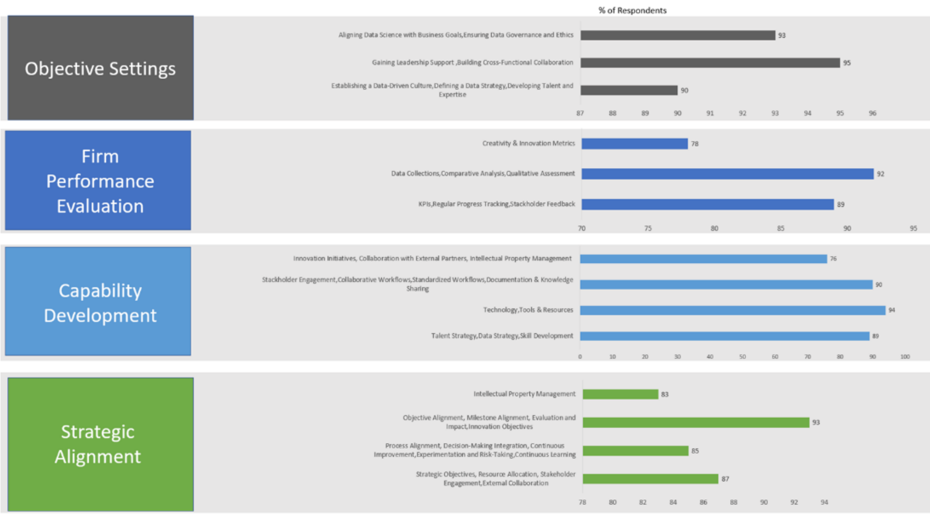

Figure 3. Thematic presentation of ecosystem development influencing factors

The participants in the study repeatedly mentioned these dimensions. Setting up business objectives was highlighted as the cycle's starting point, but many organizations were lacking in this area.

Business objectives centered on creating transparency, enabling experimentation for needs discovery, variability exploration, and performance improvement, segmenting the population for customization, and replacing/supporting human decision-making. These objectives were spread across various organizational levels, including business, process, activity, development, and innovation.

Firm performance evaluation involves developing key performance indicators (KPIs) from the business objectives, setting targets, and identifying gaps. For instance, the business objective of churn prediction in telecom can have the churn rate as a KPI, which can be shared at various organizational levels, as mentioned above. The organization must develop capabilities such as data collection and management, reporting dashboard development, and predictive modeling (including machine learning model development and deployment) to achieve the KPI.

The insights gained from the developed capabilities lead to various levels of strategic alignment within the organization. Some participants suggested reorganizing the organizational hierarchy, while others emphasized the need for investment planning.

Figure 3 below provides a high-level representation of the ecosystem development framework for enterprise data science & research projects.

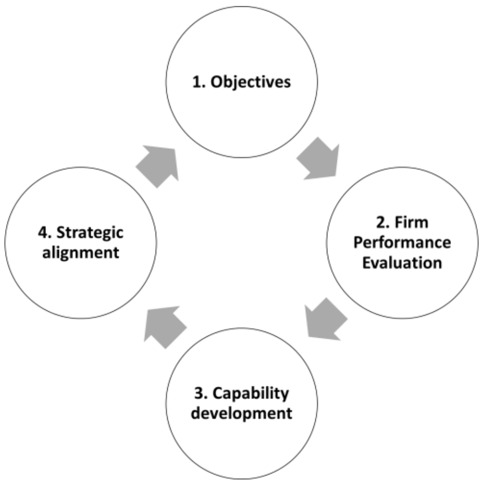

Figure 4. Enterprise Ecosystem Development Framework For Data Science & Research Projects: Broad View

The figure above represents the intricate and iterative process of the ecosystem development framework for enterprise data science and research projects through a visual model. Below is the table, Table 2, for a detailed view of the model.

Table 2.Enterprise Ecosystem Development Framework For Data Science & Research Projects: Detailed View

Steps	**1. Objectives**	**2. Firm Performance Evaluation**	**3. Capability development**	**4. Strategic alignment**		
Level				**Scope**	**Barriers to Resolve**	**Enablers to implement**
Business	• Establishing a Data-Driven Culture • Defining a Data Strategy • Gaining Leadership Support • Aligning Data Science with Business Goals • Building Cross-Functional Collaboration • Ensuring Data Governance and Ethics • Developing Talent and Expertise	KPIs, Comparative Analysis, Qualitative Assessment	Talent Strategy, Data Strategy, Technology, Stakeholder Engagement	Strategic Objectives, Resource Allocation, Stakeholder Engagement,	•Lack of top-level management support and understanding of the value of data science and research projects. •Resistance to change and a traditional mindset that does not prioritize data-driven decision-making. •Insufficient allocation of resources (financial, technological, and human) for data science and research initiatives	•Strong leadership commitment to driving a data-driven culture and strategic alignment. •Clear communication of the benefits of data science and research projects to key stakeholders. •Adequate resource allocation and investment in data infrastructure, talent acquisition, and training.
Process	Targeted Focus, Efficiency and Cost Reduction, Quality Enhancement, Continuous Improvement, Performance Measurement	KPIs, Data Collection and Analysis, Regular Progress Tracking	Skills Development, Collaborative Workflows, Tools and Resources, Process Optimization, Continuous Improvement, Agile Methodologies	Process Alignment, Decision-Making Integration, Continuous Improvement,	•Resistance to change existing processes and methodologies to incorporate data-driven approaches. •Lack of standardization and documentation of data collection, analysis, and interpretation processes. •Insufficient buy-in from process owners and stakeholders to embrace data-driven process optimization.	•Leadership support for process improvement and data-driven decision-making. •Clear guidelines and training on incorporating data science and research methodologies into existing processes. •Collaboration between process owners and data science teams to identify opportunities for optimization.
Activity	Clarity and Focus, Task Prioritization. Milestone Achievement: , Accountability and Measurement, Continuous Improvement	Clear Objectives, Milestone Achievement, Task-specific KPIs,	Skill Enhancement, Standardized Workflows, Documentation and Knowledge Sharing	Objective Alignment, Milestone Alignment, Evaluation and Impact	•Limited availability of skilled data scientists and researchers with expertise in the specific project domain. •Lack of access to high-quality data and appropriate tools for data analysis and modeling. •Insufficient time and resources allocated for conducting thorough research and data analysis.	•Talent acquisition and development programs to build a skilled workforce with domain knowledge and data science expertise. •Investment in advanced data analytics tools, technologies, and infrastructure. •Adequate resource allocation and time management to ensure thorough research and data analysis.
Development & Innovation	Encouraging Exploration, Stimulating Creativity, Risk-Taking and Learning, Collaborative Innovation, Out-of-the-Box Thinking	Creativity and Innovation Metrics, Learning and Adaptation, Stakeholder Feedback,	Innovation Initiatives, Collaboration with External Partners, Intellectual Property Management	Innovation Objectives, Experimentation and Risk-Taking: , External Collaboration, Intellectual Property Management, Continuous Learning	•Lack of a supportive organizational culture that encourages innovation, experimentation, and risk-taking. •Insufficient collaboration with external partners, limiting exposure to new methodologies and technologies. •Inadequate mechanisms for protecting and commercializing intellectual property generated through data science and research projects.	•Cultivating an innovation-friendly culture that values experimentation and rewards risk-taking. •Establishing collaborations with academic institutions, research organizations, or industry partners for knowledge exchange and access to cutting-edge advancements. •Developing intellectual property management frameworks to protect and leverage innovative solutions and algorithms.

The model portrays the interrelationships between various categories and subcategories of findings obtained from extensive interviews and feedback sessions with a sample of experts. This model was developed to address the following research question: What steps does an organization take to develop an ecosystem for data science and research projects? More granular follow-up questions follow this open-ended question.

The semi-structured interview design resulted in obtaining profound insights from the participants. During the initial interviews, a substantial amount of the first hour was left open-ended, allowing the participants to elaborate on how their organizations are fostering an ecosystem for data science and research projects.

An integrated perspective acknowledges starting points and pathways through the ecosystem development process, which supports the study's propositions and identifies categories and subcategories.

From here onward, we will explore the propositions out of which the categories and subcategories developed. The propositions are as follows.

Proposition 1: Defining Organizational Objectives

According to research findings, objective setting plays a critical role in data science and research projects at various levels within business units and functional departments during ecosystem development. It provides a distinct direction, focus, and purpose for these projects, ensuring their alignment with the overall goals and objectives of the organization. One participant emphasized the significance of defining objectives and formulating hypotheses based on them. Another participant mentioned that teams explore diverse solutions based on the set objectives. Business objectives may vary depending on the organization's goals, as indicated by one participant, who cited examples such as revenue increase, profit increase, and market expansion.

Participants also stressed the need to commence with business objectives and suggested that data exploration and dashboard transparency could aid in achieving these objectives. They added that data science and analytics could be utilized to comprehend customer behavior and expand the market.

Here are some quotes from the participants:

Quote 1: "The most crucial action for initiating any project or even introducing data science and research projects is setting the business objectives."

Quote 2: "Setting business objectives and defining problems serves as the starting point for the ecosystem."

Quote 3: "Therefore, it is necessary to establish the objective first, followed by formulating hypotheses and determining the subsequent actions."

Quote 4: "Different teams explore various solutions, depending on the objectives set."

Quote 5: "Increasing revenue is one aspect. Consequently, increased profit is likely to follow. Thus, it constitutes the primary objective."

Quote 6: "Setting business objectives must be the initial step in ecosystem development."

Quote 7: "We start with business objectives, which can encompass multiple goals. This may involve data exploration or ensuring transparency through dashboards. When you create a dashboard, people can access real data, and the organization can have a unified source of information. So, the process begins with setting business objectives or defining problem statements."

Quote 8: "Furthermore, by employing data science and analytics, we can gain insights into customer behavior, needs, and requirements, enabling us to expand our market reach. Market expansion is also one of these goals."

Quote 10: "Aligning organizational objectives is the first step towards strategic success, as it sets the direction for growth and guides decision-making."

Quote 11: "Defining clear organizational objectives is essential to ensure that all efforts are focused and coordinated towards a common purpose."

Quote 12: "Organizational objectives serve as a compass, guiding the actions and priorities of the entire organization and enabling strategic alignment."

Quote 13: "The process of defining organizational objectives lays the foundation for strategic planning and provides a framework for measuring progress and success."

Quote 15: "By clearly articulating organizational objectives, leaders empower their teams to work towards a shared vision and drive strategic alignment across the organization."

The research concluded that objectives are the fundamental driving force of the framework and must be explicitly defined and categorized. The study identified four different levels at which objectives should be defined.

A. Business Level

Business-level objective setting is a crucial step in the development and success of any organization. It serves as a guiding framework that outlines the direction, focus, and purpose of data science and research projects within different business units and functional departments. By clearly defining and categorizing objectives, organizations can ensure that these projects align with their overall goals and strategic vision. In this context, several participants in the research have emphasized the significance of objective setting and its impact on the effectiveness and outcomes of data-driven initiatives.

a. Establishing a Data-Driven Culture:

The organization aims to cultivate a culture that values data-driven decision-making and actively encourages the utilization of data science and research to drive business outcomes. This objective involves creating awareness among stakeholders, providing education on the benefits and value of data-driven insights, and fostering a mindset that embraces data-driven decision-making throughout the organization.

b. Defining a Data Strategy:

The organization seeks to develop a comprehensive data strategy aligning with its strategic objectives. This strategy encompasses the organization's vision for effectively utilizing data, including aspects such as data acquisition, storage, integration, analysis, and governance. By defining a clear data strategy, the organization aims to maximize the potential of its data assets and leverage them strategically to achieve its goals.

c. Gaining Leadership Support:

The organization endeavors to secure support from senior management and key stakeholders for its data science and research initiatives. This objective entails obtaining buy-in from leadership, allocating necessary resources, and prioritizing and including data-related initiatives in the organization's strategic agenda. By gaining leadership support, the organization can foster an environment that values and supports data-driven initiatives.

d. Aligning Data Science with Business Goals:

The organization strives to align its data science and research projects with its specific business goals and objectives. Doing so aims to identify areas where data-driven insights can have the most significant impact, such as improving operational efficiency, enhancing the customer experience, or identifying new revenue streams. This objective ensures that data initiatives are closely tied to the organization's overarching goals, thereby maximizing their effectiveness and contribution to success.

e. Building Cross-Functional Collaboration:

The organization aims to foster collaboration and interdisciplinary teamwork between its data science teams and various functional departments. By encouraging cross-functional collaboration, the organization ensures that data science and research projects effectively address the specific needs and challenges of different business functions. This objective promotes a holistic approach to leveraging data insights, drawing on diverse perspectives and expertise to drive innovation and achieve optimal results.

f. Ensuring Data Governance and Ethics:

The organization prioritizes the establishment of robust data governance frameworks and ethical guidelines to ensure the responsible and ethical use of data. This objective involves addressing key considerations such as data privacy, security, compliance, and transparency. By implementing stringent data governance practices, the organization aims to build trust within the organization and with external stakeholders, fostering confidence in the handling and use of data.

g. Developing Talent and Expertise:

The organization recognizes the importance of attracting, developing, and retaining talent with expertise in data science and research. This objective involves investing in training programs, establishing career paths, and creating opportunities for continuous learning and development in data-related fields. By nurturing talent and expertise, the organization can build a strong foundation of skilled professionals who can effectively leverage data science and research to drive innovation and success.

Some Quotes from Participants:

"The key objective is to understand business-level objectives clearly."

"Data science will help you only when you are very clear about the objective or the imperative of why you want to do the data science."

"Departmental goals must be aligned with high-level business goals."

"Shared objectives can bring collaboration and coordination among teams, preventing failures of data science and research projects."

"At the business level, strategic alignment ensures that all aspects of the organization are directed towards achieving common objectives and maximizing performance."

"Strategic alignment at the business level involves harmonizing goals, resources, and actions across different functions and departments to drive collective success."

"Business-level strategic alignment is about integrating data-driven decision-making and leveraging resources effectively to pursue shared objectives and gain a competitive edge."

"By fostering strategic alignment at the business level, organizations can enhance collaboration, break down silos, and optimize their operations to deliver superior results."

"Business-level strategic alignment enables organizations to adapt to changing market dynamics, align their activities with customer needs, and drive sustainable growth."

In summary, at the business level, the organization focuses on establishing a data-driven culture, defining a comprehensive data strategy, gaining leadership support, aligning data science with business goals, fostering cross-functional collaboration, ensuring data governance and ethics, and developing talent and expertise. These objectives aim to create an environment that values and utilizes data effectively to drive business success.

B. Process Level

Setting objectives at the process level, both within data science and research projects, as well as at the broader business process level, is a crucial step in driving improvement and achieving organizational goals. This systematic approach brings numerous benefits to the organization, as highlighted by the participants:

a. Targeted Focus:

Clear process-level objectives provide a focused direction for improvement efforts. Participants stressed the importance of understanding process-level objectives: "The key objective is to have a clear understanding of process-level objectives." By setting specific objectives, organizations can focus on areas that require enhancement, whether streamlining data collection, optimizing production workflows, or improving customer service processes.

b. Efficiency and Cost Reduction:

Objectives at the process level play a vital role in identifying inefficiencies and bottlenecks, leading to improved efficiency and cost reduction. Organizations can optimize their operations by setting objectives to streamline processes, automating repetitive tasks, and eliminating unnecessary steps. One of the participants highlighted this: "Dividing objectives at the process level makes the project doable and deliverable."

c. Quality Enhancement:

Process-level objectives can focus on enhancing the quality and reliability of outcomes. By setting objectives to improve data validation techniques, implement robust quality control measures, or enhance product/service quality, organizations ensure that their processes generate outputs of higher quality. A participant said, "Shared objectives can bring collaboration and coordination among teams and enhance quality."

d. Continuous Improvement:

The objective set at the process level fosters a culture of continuous improvement. By establishing objectives that encourage ongoing monitoring, evaluation, and refinement of processes, organizations create a feedback loop for iterative enhancements. This allows for constant learning, optimization, and adaptation to changing requirements and emerging best practices. A participant emphasized the importance of continuous improvement, stating, "It's more of a culture of continuous improvement."

e. Performance Measurement:

Process-level objectives provide measurable targets for evaluating progress and measuring the effectiveness of implemented changes. By defining key performance indicators (KPIs) aligned with process improvement objectives, organizations can track their performance, identify areas for further attention, and make data-driven decisions. A participant emphasized the need for alignment, stating, "Business level objective settings followed by activity and process level need to align for achieving organizational goals."

Here are some more quotes from participants:

"Defining clear objectives at the process level is essential for aligning data-driven initiatives with organizational goals and ensuring focused and impactful outcomes."

"Process-level objective definition establishes a roadmap for data-driven process optimization, guiding teams towards desired outcomes and performance improvements."

"By defining objectives at the process level, organizations can set measurable targets, track progress, and align data-driven efforts with key performance indicators."

"Process-level objective definition facilitates effective resource allocation, enabling organizations to prioritize initiatives that align with strategic goals and deliver maximum value."

"Clear process-level objectives provide a shared understanding among stakeholders, fostering collaboration and ensuring a unified focus on leveraging data to drive process excellence."

Setting objectives at the process level within data science and research projects and across broader business processes ensures a systematic and targeted approach to driving improvement, efficiency, and quality. It fosters a culture of continuous improvement and provides a framework for performance measurement, as supported by the participants' insights.

C. Activity Level

The objective set at the activity level is crucial for defining specific goals and outcomes for individual tasks or activities within both data science and business processes. This comprehensive approach ensures alignment and effectiveness in achieving desired outcomes. Let's explore why activity-level objectives are essential:

a. Clarity and Focus:

"Setting objectives at the activity level provides clarity and focus on the desired outcomes of each task. It ensures that team members understand what needs to be accomplished and how it contributes to the larger project goals," said one participant. Clear objectives eliminate ambiguity and enable efficient execution of activities, whether they are related to data science projects or business processes.

b. Task Prioritization:

According to a participant, "Businesses typically create their annual plan for the upcoming year, where they determine their key objectives, KPIs, and targets. Once these are finalized, they hold sessions with level one and two leaders to brainstorm how data science can align with their strategy and scorecard." Activity-level objectives help prioritize tasks within data science and business projects, ensuring that critical activities receive appropriate attention and resource allocation.

c. Milestone Achievement:

"Objective setting at the activity level helps in defining milestones or checkpoints that mark significant progress in the project," explained one participant. These milestones are tied to specific objectives, allowing teams to track their progress and ensure alignment with the desired outcomes within the project timeline, whether related to data science initiatives or business process improvements.

d. Accountability and Measurement:

"Clear objectives provide a basis for measuring performance and holding team members accountable," emphasized a participant. Well-defined objectives make it easier to assess the effective execution of activities and the achievement of desired outcomes, whether they are within the realm of data science or business processes. This promotes timely course correction and ensures project alignment.

e. Continuous Improvement:

A participant highlighted that "activities like sales assessment, cost-cutting for back office, etc., are good to define as objectives in advance." The objective set at the activity level promotes a culture of continuous improvement within both data science and business projects. By defining objectives that encourage learning, experimentation, and the adoption of innovative approaches, teams can continually enhance their methods and techniques in both domains.

Here are quotes from participants:

"Defining objectives at the activity level sets the direction for data-driven projects, ensuring a focused and purposeful approach to data analysis and research."

"Activity-level objective definition provides a roadmap for data scientists and researchers, guiding their efforts to extract insights, build models, and generate impactful results."

"By clearly defining objectives at the activity level, organizations can align data science projects with business needs, ensuring that resources and efforts are directed towards the most valuable activities."

"Activity-level objective definition facilitates effective resource allocation, allowing organizations to allocate the right talent, tools, and time to each data-driven activity for optimal outcomes."

"Clear objectives at the activity level enhance project management and decision-making, enabling stakeholders to monitor progress, evaluate performance, and make informed adjustments along the data science journey."

Setting objectives at the activity level ensures that tasks and activities within data science and business projects align with the overall goals and objectives. It provides clarity, helps prioritize tasks, facilitates milestone achievement, fosters accountability and measurement, and promotes continuous improvement.

D. Development & Innovation Level

The objective set at the development and innovation level plays a pivotal role in fostering creativity, exploration, and the generation of novel solutions within data science and research projects. This level focuses on pushing boundaries, embracing risk-taking, and promoting out-of-the-box thinking. Let's explore the reasons why it is crucial, with insights from different participants:

A participant states, "Development and innovation level objective setting can give new opportunities to expand the organizational business." This level provides the opportunity to develop new products, services, or innovate existing processes, ultimately driving growth and expansion.

a. Encouraging Exploration:

A participant shares his perspective: "Objectives at the development and innovation level promote the exploration of new ideas, methods, and technologies. It encourages experimentation and pursuing innovative approaches, driving creativity and pushing the boundaries within data science and research projects."

b. Stimulating Creativity:

Participants emphasize the importance of objectives that foster creativity, stating, "Objectives emphasizing development and innovation inspire team members to think outside the box and develop creative solutions. This creates a supportive environment where new ideas are encouraged and diverse perspectives are valued, leading to breakthrough discoveries and unique insights."

c. Risk-Taking and Learning:

A participant discusses the importance of risk-taking and learning at this level, noting, "Objectives at the development and innovation level often involve taking calculated risks and embracing a learning mindset. By encouraging experimentation, organizations create a safe space for team members to take risks, learn from failures, and iterate on their approaches. This induces a culture of innovation and continuous learning within the project team."

d. Collaborative Innovation:

Participants highlighted the benefits of collaboration in innovation. A participant said, "Development and innovation level objectives can focus on promoting cross-functional collaboration and knowledge sharing. By leveraging the diverse expertise of team members from different disciplines, organizations can develop interdisciplinary approaches and innovative solutions that would not have been possible in siloed environments."

e. **Out-of-the-Box Thinking:**

A participant shares his experience: "Objectives at this level challenge conventional thinking and push for unconventional solutions. By encouraging team members to think differently and question assumptions, organizations uncover new insights, hidden patterns, and arrive at innovative solutions that give them a competitive advantage."

Setting objectives at the development and innovation level within data science and research projects fosters creativity, exploration, and the generation of novel solutions. It stimulates the pursuit of new ideas, encourages risk-taking and learning, promotes collaborative innovation, and encourages out-of-the-box thinking. By setting clear objectives at this level, organizations foster a culture of innovation, drive breakthrough discoveries, and stay at the forefront of data-driven advancements in their field.

Proposition 2: Firm Performance Evaluation

Organizations can utilize several essential elements and methods when evaluating firm performance concerning the abovementioned objectives. Here are key considerations at each level of evaluation:

A. **Business Level Evaluation:**

At the business level, evaluating performance is crucial to assess the organization's overall strategic alignment and success. Key considerations include:

a. **KPIs**:

Define key performance indicators (KPIs) that align with the strategic objectives of the business unit. These metrics may include revenue growth, market share, profitability, customer retention, or brand perception.

b. **Comparative Analysis**:

Comparing the business's performance against industry benchmarks and competitors provides valuable insights into its market position and areas for improvement. This analysis helps identify strengths and weaknesses. A participant emphasized the importance of benchmarking: "Performance evaluation is a process that involves error minimization and is crucial for achieving business objectives. The evaluation process involves using various KPIs to assess the efficiency and effectiveness of different activities."

c. Qualitative Assessment:

Gathering qualitative data through customer surveys, market research, and stakeholder feedback provides valuable insights into the business's products, services, and brand perceptions. This feedback helps understand customer satisfaction and brand reputation. A participant mentioned, "We further look at our business objectives and qualitatively assess the current situation."

Here are some quotes from participants:

"Firm performance evaluation at the business level provides a holistic view of the organization's overall success and guides strategic decision-making for future growth and sustainability."

"Evaluating firm performance at the business level helps identify strengths, weaknesses, and areas of improvement, enabling proactive measures to enhance competitiveness and achieve strategic objectives."

"By conducting thorough performance evaluation at the business level, organizations can assess the effectiveness of their strategies, investments, and resource allocation, ensuring alignment with long-term goals."

Evaluating the business level performance allows organizations to align their strategies, measure their success against industry benchmarks, and gain a deeper understanding of customer perceptions and satisfaction. It enables them to identify areas for improvement and make informed decisions to drive overall business growth and success.

B. **Process Level Evaluation**:

At the process level, evaluating performance involves assessing the efficiency and effectiveness of the organization's systems, workflows, and processes. Key considerations include:

a. **KPIs**:

Define KPIs related to process efficiency, cost reduction, or quality improvement. Examples of process-related KPIs include cycle time, error rates, cost per unit, or customer satisfaction with the process. These KPIs help measure the performance of specific processes and identify areas for improvement.

A participant emphasized the importance of KPIs at the process level, stating, "In our organization, we define specific KPIs for each process to ensure that we are effectively measuring efficiency, cost, and quality. These KPIs provide us with valuable insights into process performance and help us identify areas for optimization."

b. Data Collection and Analysis:

Collect data on process performance and apply analytical techniques to identify bottlenecks, inefficiencies, or areas for optimization. This may involve analyzing process metrics, such as throughput, lead time, or resource utilization, to gain insights into the effectiveness of processes and identify areas that require attention.

A participant highlighted the significance of data collection and analysis in process evaluation, stating, "We collect and analyze data on various process metrics to understand how well our processes are performing. By using analytical techniques, we can identify bottlenecks, inefficiencies, and opportunities for optimization."

c. Regular Progress Tracking:

Monitor process metrics over time to track progress, identify trends, and implement necessary adjustments for continuous improvement. By tracking process performance regularly, organizations can identify deviations from desired outcomes, take corrective actions, and drive ongoing enhancements.

A participant emphasized the importance of regular progress tracking: "We have implemented a system where we regularly track process metrics and monitor their performance over time. This helps us identify trends, spot areas that need improvement, and make necessary adjustments for continuous improvement."

Process level evaluation allows organizations to optimize their operational workflows, improve efficiency, and enhance the quality of their outputs. It provides insights into the effectiveness of specific processes, highlights areas for improvement, and enables organizations to implement targeted interventions to streamline operations and enhance overall process performance.

A participant emphasized the significance of process-level evaluation, stating, "Evaluating performance at the process level is crucial as it allows us to assess the efficiency and effectiveness of our operational workflows. By identifying areas for improvement and implementing targeted interventions, we can enhance our overall process performance and drive continuous improvement.".

Here are quotes from participants:

"Firm performance evaluation at the process level focuses on assessing the effectiveness and efficiency of key operational processes, enabling process optimization and continuous improvement."

"Evaluating performance at the process level helps identify bottlenecks, inefficiencies, and areas for automation or streamlining, leading to enhanced productivity and cost-effectiveness."

"By monitoring and evaluating firm performance at the process level, organizations can ensure that processes align with strategic objectives, customer expectations, and industry best practices."

C. Activity Level Evaluation:

Activity level evaluation focuses on assessing the performance of individual tasks or activities within a project. This evaluation helps ensure that activities align with project goals and objectives. Critical considerations for activity level evaluation include:

a. Clear Objectives:

Establish clear objectives for each activity within the project and align them with the overall project goals. Clear objectives provide a roadmap for team members, guiding their efforts and ensuring they understand each activity's desired outcomes.

A participant highlighted the importance of clear objectives at the activity level: "When we define objectives for each activity, it helps us align our efforts with the overall project goals. Clear objectives give us a sense of direction and clarity on what needs to be achieved."

b. Milestone Achievement:

Track and evaluate the achievement of milestones and objectives at each activity level to gauge progress and ensure alignment with the project's objectives. Milestones serve as checkpoints, marking significant progress and helping to measure the overall advancement of the project.

A participant emphasized the significance of milestone achievement in activity level evaluation, stating, "We track milestones and objectives at each activity level to ensure that we are progressing in line with our project's goals. Achieving milestones helps us gauge our progress and stay on track."

c. Task-specific KPIs:

Define task-specific KPIs to evaluate the effectiveness and efficiency of individual activities. These KPIs can include metrics such as completion time, error rates, adherence to quality standards, or any other relevant performance indicators specific to each activity.

A participant discussed the importance of task-specific KPIs, stating, "For each activity, we define KPIs that allow us to evaluate its effectiveness and efficiency. These KPIs help us measure the performance of individual tasks and identify areas where improvements can be made."

Activity level evaluation enables organizations to monitor the progress of individual tasks, ensure alignment with project objectives, and identify areas for improvement. It provides insights into the effectiveness and efficiency of activities, helping teams stay on track and make necessary adjustments to achieve project success.

A participant highlighted the significance of activity level evaluation: "Evaluating performance at the activity level is crucial for ensuring that our tasks are aligned with the project goals. It allows us to monitor progress, identify areas of improvement, and make necessary adjustments to stay on track."

Here are some more quotes from participants:

"Performance evaluation at the activity level enables organizations to assess the impact and value generated by specific data-driven activities, guiding resource allocation and decision-making."

"Evaluating firm performance at the activity level provides insights into the effectiveness of data analysis, research methodologies, and project execution, facilitating continuous learning and optimization."

"By measuring and evaluating performance at the activity level, organizations can identify high-performing activities, replicate successful approaches, and address any gaps or challenges hindering performance."

By conducting thorough activity-level evaluations, organizations can enhance task execution, improve project outcomes, and ensure that each activity contributes effectively to the overall project's success.

D. **Development & Innovation Level Evaluation**:

Development and innovation level evaluation focus on assessing the organization's ability to foster creativity, explore, and generate novel solutions within data science and research projects. This evaluation aims to measure the effectiveness of development and innovation efforts and their impact on overall performance. Key considerations for development and innovation level evaluation include:

a. **Creativity and Innovation Metrics**:

Define metrics that measure the outcomes of development and innovation efforts. These metrics can include the number of new ideas generated, the successful implementation of innovative solutions, the impact of innovations on key performance indicators, or any other relevant measures that capture the organization's ability to drive creativity and innovation.

A participant emphasized the importance of creativity and innovation metrics: "In our evaluation, we look at metrics that reflect the outcomes of our development and innovation efforts. This includes measuring the number of new ideas generated and the successful implementation of innovative solutions. These metrics help us assess our organization's ability to drive creativity and innovation."

b. **Learning and Adaptation**:

Evaluate the organization's ability to learn from failures, iterate on approaches, and adapt based on feedback to drive continuous improvement and innovation. This includes assessing the organization's learning culture, willingness to take calculated risks, and capacity to embrace change and adapt to evolving market conditions.

A participant discussed the significance of learning and adaptation in development and innovation level evaluation, stating, "Our evaluation includes assessing our organization's ability to learn from failures and adapt to changes. We need a culture of continuous improvement and a willingness to embrace new approaches and ideas."

c. Stakeholder Feedback:

Gather stakeholders' feedback to assess the impact and effectiveness of the development and innovation initiatives. Stakeholder feedback provides valuable insights into how the organization's development and innovation efforts are perceived, their contribution to organizational goals, and areas where improvements can be made.

A participant highlighted the importance of stakeholder feedback: "We actively seek feedback from stakeholders to understand how our development and innovation initiatives are perceived. Their input helps us assess the effectiveness of our efforts and identify areas where we can further enhance our impact."

Development and innovation level evaluation allows organizations to gauge the effectiveness of their efforts to foster creativity, exploration, and the generation of novel solutions. It helps measure the organization's ability to adapt and learn and the impact of innovations on overall performance. By conducting comprehensive evaluations at this level, organizations can drive continuous improvement, enhance innovation capabilities, and stay at the forefront of their field.

A participant emphasized the importance of development and innovation level evaluation, stating, "Evaluating our organization's development and innovation efforts is crucial for understanding our current state and determining how far we need to go to achieve our business goals. It enables us to assess our innovation capabilities and make necessary adjustments to drive future success."

Some more quotes from participants:

"Evaluating development and innovation at the organizational level enables organizations to gauge their overall innovative capacity, track progress towards strategic goals, and drive continuous improvement in their innovation capabilities."

"Performance evaluation at the collaborative level assesses the effectiveness of collaborative partnerships, joint ventures, and open innovation initiatives, enabling organizations to measure the value generated through collaborative efforts."

"Evaluating development and innovation at the collaborative level provides insights into the success of collaborative projects, identifies opportunities for enhanced collaboration, and facilitates knowledge sharing and co-creation of innovative solutions."

By effectively evaluating development and innovation efforts, organizations can foster a culture of innovation, drive breakthrough discoveries, and stay competitive in an ever-evolving landscape.

Organizations can comprehensively evaluate firm performance by applying these elements and methods at each level. This multi-dimensional assessment considers strategic alignment, operational efficiency, customer satisfaction, employee engagement, innovation outputs, and the overall achievement of objectives at different levels within the organization.

Proposition 3: Capability development

In today's rapidly evolving business landscape, organizations recognize the critical need to develop capabilities that enable them to effectively leverage data and analytics for strategic decision-making and sustainable growth. Capability development involves the process of acquiring and enhancing the necessary skills, resources, and organizational infrastructure to harness the power of data science, research, and business acumen. Here is the various level of capability development that can be needed after firm performance evaluation.

A. Business Level:

At the Business Level, capability development involves enhancing the organization's strategic and operational capabilities to effectively leverage data science, research, and business acumen. This level focuses on aligning talent strategy, data strategy, technology infrastructure, and stakeholder engagement to drive data-driven decision-making and foster a culture of innovation.

a. Talent Strategy:

Developing a talent strategy is crucial for building the necessary capabilities at the business level. This strategy aims to identify the skills and expertise required at the leadership and managerial levels to drive data-driven decision-making and promote a culture of innovation. It involves assessing the organization's current talent pool, identifying skill gaps, and implementing strategies to acquire, develop, and retain the right talent. A quote from the participant highlights the importance of talent capability, stating, "Technical knowledge and subject knowledge are very important for BDA. They contribute to the end goal."

b. Data Strategy:

A comprehensive data strategy is essential for effective capability development at the business level. This strategy defines the organization's data governance practices, acquisition methods, and integration approaches. It ensures data quality, security, and compliance while facilitating accessibility and usability for data science and research projects. The participant emphasized the significance of planning and investment, connectivity, and compatibility, stating, "Planning and investment, technologically, connectivity, compatibility, investments, and all those things." This quote highlights the need to define a robust data strategy that aligns with the organization's objectives.

c. Technology Infrastructure:

Investing in robust and scalable technology infrastructure is critical for capability development at the business level. This includes infrastructure for storing, processing, and analyzing big data and adopting advanced analytics tools and technologies. Machine learning, predictive modeling, natural language processing, data integration, cleansing, and visualization tools are examples of technologies that enable effective data analysis. The participant mentioned the importance of connectivity among data sources, stating, "Connectivity among the data is crucial. Silos should be avoided." This quote emphasizes the need for organizations to invest in technologies that ensure the seamless integration of diverse data sources and provide intuitive visualizations for effective data exploration and interpretation.

d. Stakeholder Engagement:

Engaging senior management and key stakeholders is vital for successful capability development at the business level. It involves communicating the strategic importance of data-driven insights and securing their support and involvement in data science and research initiatives. This induces a culture of innovation and ensures that the organization's objectives align with data-driven decision-making. Effective stakeholder engagement facilitates resource allocation, decision-making, and the implementation of data-driven strategies.

Here are some quotes from participants:

"Capability development at the business level is crucial for fostering a culture of innovation, agility, and adaptability, enabling organizations to stay ahead of market trends, seize new opportunities, and drive sustainable growth."

"Investing in capability development at the business level empowers organizations to build diverse competencies, align strategic goals with talent development, and enhance overall organizational performance."

"By developing capabilities at the business level, organizations can cultivate a dynamic workforce, nurture leadership potential, and create a strong foundation for achieving long-term success."

By addressing these elements at the business level, organizations can develop the necessary capabilities to harness the power of data science, research, and business acumen. This enables them to make informed decisions, drive innovation, and achieve sustainable growth.

B. Process Level:

At the Process Level, capability development focuses on integrating data science and research methods into existing organizational processes to optimize efficiency, enhance decision-making, and drive continuous improvement. This level involves analyzing and improving processes, implementing best practices, and adopting agile methodologies to effectively utilize data science and research capabilities.

a. Process Optimization:

Process optimization is a key aspect of capability development at the process level. It involves analyzing existing processes and identifying opportunities to integrate data science and research methods for improved efficiency and decision-making. Organizations must establish the best data collection, analysis, and interpretation practices and incorporate them into standard operating procedures. By optimizing processes, organizations can enhance their ability to collect, analyze, and utilize data effectively, leading to more informed decision-making and improved outcomes.

b. Continuous Improvement:

Capability development at the process level requires a culture of continuous improvement. Organizations should regularly review and refine data science and research processes to ensure they align with evolving business needs and technological advancements. Implementing feedback loops, monitoring performance metrics, and leveraging insights gained from data analysis are essential for driving process enhancements. The participant mentioned the importance of continuous improvement, stating, "Regularly reviewing and refining data science and research processes is crucial."

c. Agile Methodologies:

Adopting agile methodologies, such as Scrum or Kanban, is beneficial for capability development at the process level. Agile methodologies promote iterative development, collaboration, and flexibility, enabling faster delivery of insights and solutions within data science and research projects. By embracing agile methodologies, organizations can respond quickly to changing requirements, adapt their processes accordingly, and foster effective teamwork between data science/research teams and functional teams.

Here are some quotes from participants:

"Capability development at the process level enhances operational efficiency, streamlines workflows, and drives continuous improvement, enabling organizations to deliver high-quality products and services to customers."

"Investing in process-level capability development enables organizations to optimize key business processes, leverage data-driven insights, and proactively identify opportunities for innovation and efficiency gains."

"By developing capabilities at the process level, organizations can foster a culture of operational excellence, standardize best practices, and create a framework for sustained process improvement."

In summary, capability development at the process level involves optimizing existing processes, embracing continuous improvement, and adopting agile methodologies. By integrating data science and research methods into organizational processes, organizations can enhance efficiency, decision-making, and collaboration, resulting in improved outcomes and a competitive advantage in the market.

C. Activity Level:

At the Activity Level, capability development focuses on enhancing the skills, workflows, and documentation practices related to data science and research activities. It aims to maximize the effectiveness and efficiency of individual activities within data science projects. Here are key aspects of capability development at the activity level:

a. Skill Enhancement:

Capability development at the activity level emphasizes providing specific training and development opportunities to individuals involved in data science and research activities. This includes enhancing their technical skills in areas such as advanced statistical analysis, machine learning techniques, programming languages, and data visualization. By investing in skill enhancement, organizations can ensure that their team members have the necessary expertise to effectively perform their roles and contribute to the success of data science projects.

b. Standardized Workflows:

Establishing standardized workflows and protocols is crucial for capability development at the activity level. These workflows outline the steps and procedures for data collection, preprocessing, modeling, and evaluation. By standardizing workflows, organizations ensure consistency, reduce errors, and facilitate efficient collaboration among team members working on different activities within data science and research projects. The participant highlighted the importance of standardized workflows: "Establishing standardized workflows and protocols is essential for consistency and efficient collaboration."

c. Documentation and Knowledge Sharing:

Encouraging documentation and knowledge-sharing practices within project teams is a vital aspect of capability development at the activity level. This includes maintaining comprehensive methodologies, algorithms, code repositories, and project insights documentation. Effective knowledge sharing enables team members to leverage each other's expertise, promotes the reusability of solutions, and facilitates smoother transitions between team members. The participant emphasized the significance of documentation and knowledge sharing: "Comprehensive documentation of methodologies, algorithms, and project insights facilitate knowledge transfer and reusability."

Some more quotes from the participants:

"Capability development at the activity level empowers employees to excel in their specific roles, acquire specialized skills, and contribute to the overall performance and success of the organization."

"Investing in activity-level capability development enables organizations to build a skilled workforce, promote job satisfaction and employee engagement, and enhance productivity and effectiveness in day-to-day operations."

"By developing capabilities at the activity level, organizations can unlock individual potential, foster a culture of continuous learning, and drive innovation and creativity within teams."

By focusing on skill enhancement, standardized workflows, and documentation and knowledge sharing, organizations can enhance their capability at the activity level. This, in turn, leads to improved coordination, productivity, and quality within data science and research activities. It enables teams to work more efficiently, make informed decisions, and deliver impactful outcomes.

D. Development & Innovation Level

At the Development & Innovation Level, capability development aims to foster a culture of innovation and enable organizations to leverage data science and research for transformative advancements. Here are key aspects of capability development at the development and innovation level:

a. Innovation Initiatives:

Capability development at this level involves encouraging and supporting innovation within data science and research projects. Organizations should provide resources and create an environment that allows for exploring new methodologies, tools, or approaches. By including innovation, organizations can push the boundaries of what is possible and uncover novel solutions to complex problems. The participant emphasized the importance of innovation, stating, "Encouraging innovation within data science and research projects is essential for transformative advancements."

b. Collaboration with External Partners:

Collaborating with academic institutions, research organizations, or industry partners is crucial for capability development at the development and innovation level. These external partnerships bring fresh perspectives, domain expertise, and access to cutting-edge research. By leveraging external expertise, organizations can stay at the forefront of data science and research advancements, engage in joint research projects, and participate in collaborative innovation initiatives. The participant mentioned the significance of collaboration with external partners: "Fostering collaborations with external partners allows organizations to leverage external expertise and drive innovation."

c. Intellectual Property Management:

Developing procedures for protecting intellectual property generated through data science and research projects is an essential aspect of capability development at this level. Organizations must establish guidelines for patenting, licensing, or commercializing innovative solutions or algorithms developed within the organization. By effectively managing intellectual property, organizations can capitalize on their innovative ideas, gain a competitive advantage, and drive value from their data science and research efforts.

Here are some more quotes from participants:

"Capability development at the development and innovation level is instrumental in driving breakthrough ideas, harnessing emerging technologies, and creating a culture of experimentation and entrepreneurial mindset."

"Investing in capability development at the development and innovation level empowers organizations to stay at the forefront of industry trends, disrupt traditional business models, and create innovative solutions that address evolving customer needs."

"By developing capabilities at the development and innovation level, organizations can foster a culture of creativity, collaboration, and risk-taking, driving transformative growth and positioning themselves as industry leaders."

By emphasizing innovation initiatives, collaboration with external partners, and intellectual property management, organizations can enhance their capability at the development and innovation level. This enables them to drive meaningful innovation, leverage external expertise, and capitalize on their intellectual property effectively. Ultimately, it positions organizations to achieve breakthrough advancements and maintain a competitive edge in the ever-evolving data-driven landscape.

Proposition 4: Strategic alignment

Strategic alignment is crucial in minimizing barriers and maximizing enablers to achieve data science and research project objectives. These barriers and enablers can be classified at different levels as follows:

A. Business Level:

Strategic alignment at the Business Level is crucial for the successful execution of data science and research projects. It involves ensuring that these initiatives are closely aligned with the organization's strategic objectives, effectively allocating resources to support them, and actively engaging stakeholders throughout the process.

a. Strategic Objectives:

Strategic objectives are the foundation for strategic alignment, guiding data science and research projects to contribute directly to the organization's goals. However, several barriers can impede strategic alignment at the Business Level:

a.1 Barriers:

- Lack of top-level management support and understanding of the value of data science and research projects.
- Resistance to change and a traditional mindset that does not prioritize data-driven decision-making.
- To overcome these barriers and achieve strategic alignment, organizations must leverage the following enablers:

a.2 Enablers:

- Strong leadership commitment to driving a data-driven culture and strategic alignment.
- Clear communication of data science and research project benefits key stakeholders.

By addressing these barriers and leveraging enablers, organizations can align data science and research projects with strategic objectives, ensuring their relevance and contribution to the overall organizational goals.

b. Resource Allocation:

Effective resource allocation is essential for supporting data science and research projects aligned with strategic objectives. However, organizations often face barriers to allocating resources adequately:

b.1 Barrier:

- Insufficient allocation of resources (financial, technological, and human) for data science and research initiatives.
- To enable strategic alignment through resource allocation, organizations should consider the following enablers:

b.2 Enablers:

- Adequate resource allocation and investment in data infrastructure, talent acquisition, and training.
- Organizations can empower data science and research initiatives by ensuring sufficient resources are allocated to drive strategic alignment and achieve desired outcomes.

c. Stakeholder Engagement:

Engaging stakeholders is critical for strategic alignment at the Business Level, as it ensures their active involvement, alignment with goals, and support for data science and research projects. However, organizations face barriers to stakeholder engagement:

c.1 Barriers:

- Lack of awareness and understanding of the potential applications of data science and research within functional departments.
- To enable stakeholder engagement and enhance strategic alignment, organizations can leverage the following enablers:

c.2 Enablers:

- Clear communication of data science and research benefits and value-added within each functional area.
- A culture that promotes knowledge-sharing and interdisciplinary approaches facilitates cross-functional collaboration and engagement.

Organizations can foster stakeholder engagement by addressing barriers and leveraging enablers, ensuring that data science and research projects align with strategic objectives and receive the necessary support for successful execution.

Here are some quotes from participants:

"Strategic alignment at the business level is the key to unlocking the full potential of data science and research projects, ensuring they are closely tied to the organization's strategic objectives and garnering the necessary support from stakeholders."

"Effective strategic alignment at the business level requires clear communication of the benefits of data science and research projects, strong leadership commitment, and a culture that embraces data-driven decision-making."

"By aligning data science and research projects with strategic objectives at the business level, organizations can overcome barriers, leverage enablers, and maximize the value of these initiatives."

By considering strategic objectives, resource allocation, and stakeholder engagement within the context of strategic alignment at the Business Level, organizations can overcome barriers and leverage enablers to ensure that data science and research projects align with strategic goals, effectively utilize resources, and garner stakeholder support. This holistic approach fosters successful outcomes and maximizes the value of data-driven decision-making within the organization.

B. Process Level:

Strategic alignment at the Process Level is vital for effectively integrating data science and research methodologies into existing processes. It involves overcoming barriers and leveraging enablers to ensure that processes are aligned with data-driven approaches, standardized, and embraced by stakeholders.

a. Process Improvement:

Process improvement plays a crucial role in strategic alignment, as it involves incorporating data-driven approaches into existing processes. However, organizations may face barriers in driving process improvement:

a.1 Barriers:

- Resistance to change existing processes and methodologies to incorporate data-driven approaches.

To enable strategic alignment through process improvement, organizations should consider the following enablers:

a.2 Enablers:

- Leadership support for process improvement and data-driven decision-making.
- Clear guidelines and training on incorporating data science and research methodologies into existing processes.

Organizations can drive process improvement by addressing these barriers and leveraging enablers, ensuring that data-driven approaches are integrated seamlessly into existing processes.

b. Standardization:

Standardization ensures consistency and reliability in data collection, analysis, and interpretation processes. However, organizations may encounter barriers to achieving standardization:

b.1 Barriers:

- Lack of standardization and documentation of data collection, analysis, and interpretation processes.

To enable strategic alignment through standardization, organizations can leverage the following enablers:

b.2 Enablers:

- Clear guidelines and training on standardized data collection, analysis, and interpretation processes.
- Collaboration between process owners and data science teams to identify opportunities for optimization.

By addressing barriers and leveraging enablers, organizations can establish standardized processes that align with data-driven approaches, facilitating strategic alignment and ensuring reliable outcomes.

c. Stakeholder Buy-in:

Gaining stakeholder buy-in is crucial for the successful integration of data-driven processes. However, organizations may face barriers in obtaining buy-in from process owners and stakeholders:

c.1 Barrier:

- Insufficient buy-in from process owners and stakeholders to embrace data-driven process optimization.

To enable stakeholder buy-in and enhance strategic alignment, organizations should consider the following enablers:

c.2 Enablers:

- Clear communication of the benefits and value-added of data-driven processes to process owners and stakeholders.
- Collaboration between process owners and data science teams to demonstrate the value and impact of data-driven optimizations.

Organizations can gain stakeholder buy-in by addressing barriers and leveraging enablers, ensuring that data-driven processes are embraced and strategically aligned with the organization's goals.

Quotes from the participants:

"Strategic alignment at the process level ensures that data science and research methodologies are seamlessly integrated into existing processes, driving efficiency and optimizing outcomes."

"To achieve strategic alignment at the process level, organizations must address barriers such as resistance to change and embrace enablers such as leadership support and clear guidelines on incorporating data-driven approaches."

"By aligning processes with data-driven approaches, organizations can establish standardized practices, enhance reliability, and promote strategic alignment in data science and research projects."

By considering process improvement, standardization, and stakeholder buy-in within the context of strategic alignment at the Process Level, organizations can overcome barriers and leverage enablers to ensure that data science and research methodologies are effectively integrated into existing processes. This promotes strategic alignment, facilitates data-driven decision-making, and drives overall organizational success.

C. Activity Level:

Activity Level is a critical aspect of strategic alignment, focusing on the execution of specific data science and research activities. Organizations must address barriers and leverage enablers at the Activity Level to achieve successful outcomes.

a. Skilled Workforce:

A skilled and capable workforce is essential for effectively executing data science and research activities. However, organizations may face barriers related to the availability of skilled professionals:

a.1 Barriers:

- Limited availability of skilled data scientists and researchers with expertise in the specific project domain.

To overcome this barrier and ensure strategic alignment at the Activity Level, organizations should consider the following enablers:

a.2 Enablers:

- Talent acquisition and development programs to build a skilled workforce with domain knowledge and data science expertise.

By investing in talent acquisition and development programs, organizations can build a capable workforce aligning with strategic objectives and effectively executing data science and research activities.

b. Data and Tools:

Access to high-quality data and appropriate tools is crucial for conducting thorough research and data analysis. However, organizations may face barriers to obtaining the necessary resources:

b.1 Barriers:

- Lack of access to high-quality data and appropriate data analysis and modeling tools.

To enable strategic alignment at the Activity Level, organizations can leverage the following enablers:

b.2 Enablers:

- Investment in advanced data analytics tools, technologies, and infrastructure.
- Adequate resource allocation and time management to ensure thorough research and data analysis.

By investing in data analytics tools, technologies, and infrastructure and allocating sufficient resources, organizations can effectively enhance their capability to execute data science and research activities.

c. Resource Allocation:

Proper allocation of resources, including financial, technological, and human resources, is crucial for the successful execution of data science and research projects. However, organizations may encounter barriers related to resource constraints:

c.1 Barriers:

- Insufficient time and resources were allocated for conducting thorough research and data analysis.

To ensure strategic alignment at the Activity Level, organizations should focus on the following enablers:

c.2 Enablers:

- Adequate resource allocation and investment in data infrastructure, talent acquisition, and training.

By allocating resources appropriately and investing in the necessary infrastructure, talent acquisition, and training, organizations can align their activities with strategic objectives and enhance their capacity to effectively execute data science and research projects.

Here are quotes from the participants:

"Strategic alignment at the activity level focuses on the execution of specific data science and research activities, requiring a skilled workforce, access to quality data and tools, and proper resource allocation."

"Overcoming barriers such as limited availability of skilled professionals and leveraging enablers like talent acquisition programs and adequate resource allocation is essential for strategic alignment at the activity level."

"By aligning activities with strategic objectives, organizations can execute data science and research projects effectively, driving impactful outcomes and strategic alignment."

Organizations can ensure that data science and research activities are executed efficiently and aligned with strategic objectives by addressing barriers and leveraging enablers at the Activity Level. This promotes strategic alignment throughout the project lifecycle, leading to valuable insights, informed decision-making, and overall success in achieving organizational goals.

D. Development & Innovation Level:

Development & Innovation Level plays a crucial role in strategic alignment, focusing on fostering development and innovation in data science and research projects. Organizations must address barriers and leverage enablers at the Development & Innovation Level to ensure successful outcomes.

a. Organizational Culture:

Creating a supportive organizational culture that encourages innovation, experimentation, and risk-taking is essential for fostering development and innovation in data science and research projects. However, organizations may face barriers related to their culture:

a.1 Barriers:

- Lack of a supportive organizational culture that encourages innovation, experimentation, and risk-taking.

To enable strategic alignment at the Development & Innovation Level, organizations should consider the following enablers:

a.2 Enablers:

- Cultivating an innovation-friendly culture that values experimentation and rewards risk-taking.

By fostering an innovation-friendly culture, organizations can create an environment that promotes development and innovation in data science and research projects.

b. Collaboration and Knowledge Exchange:

Collaboration with external partners, such as academic institutions, research organizations, or industry partners, is crucial for accessing new methodologies, technologies, and advancements. However, organizations may face barriers related to collaboration and knowledge exchange:

b.1 **Barriers**:

- Insufficient collaboration with external partners, limiting exposure to new methodologies and technologies.

To enable strategic alignment at the Development & Innovation Level, organizations can leverage the following enablers:

b.2 **Enablers**:

- Establishing collaborations with academic institutions, research organizations, or industry partners for knowledge exchange and access to cutting-edge advancements.

By fostering collaborations and partnerships, organizations can gain exposure to new methodologies, technologies, and advancements, fostering development and innovation in data science and research projects.

c. Intellectual Property Management:

Protecting and commercializing intellectual property generated through data science and research projects is crucial for leveraging innovative solutions and algorithms. However, organizations may encounter barriers related to intellectual property management:

c.1 Barriers:

- Inadequate mechanisms for protecting and commercializing intellectual property generated through data science and research projects.

To enable strategic alignment at the Development & Innovation Level, organizations should focus on the following enablers:

c.2 **Enablers**:

- Developing intellectual property management frameworks to protect and leverage innovative solutions and algorithms.

By implementing intellectual property management frameworks, organizations can safeguard their innovative solutions and algorithms, allowing for commercialization and maximizing their value.

Here are some quotes from the participants:

"Strategic alignment at the development and innovation level is critical for fostering a culture of development and innovation within data science and research projects."

"Organizational culture, collaboration, and intellectual property management play pivotal roles in enabling strategic alignment at the development and innovation level."

"By creating an innovation-friendly culture, fostering collaborations, and implementing intellectual property management frameworks, organizations can promote development and innovation in data science and research projects, achieving strategic alignment."

Organizations can induce development and innovation in data science and research projects by addressing barriers and leveraging enablers at the Development & Innovation Level. It promotes strategic alignment by creating an environment that encourages creativity, collaboration, and the protection and commercialization of valuable intellectual property.

By addressing these barriers and leveraging the enablers, organizations can create an environment conducive to objective setting, firm performance evaluation, capability development, and strategic alignment in data science and research projects.

4.6 Summary of Findings

The research involved a semi-structured interview to gain participants' deep insights regarding their organizations' efforts in fostering data science and research ecosystems. The interviews allowed participants to elaborate on how their organizations approach these projects. The study identified four main categories: objective setting, firm performance evaluation, capability development, and strategic alignment, along with several subcategories.

The objective setting was found to be crucial at various levels within business units and functional departments. It provides direction, focus, and purpose, ensuring alignment with the organization's overall goals. The research concluded that objectives are fundamental to the framework and should be explicitly defined and categorized.

At the business level, objectives include establishing a data-driven culture, defining a data strategy, gaining leadership support, aligning data science with business goals, fostering cross-functional collaboration, ensuring data governance and ethics, and developing talent and expertise. These objectives create an environment that values and utilizes data effectively to drive business success.

At the process level, objectives help drive improvement and achieve organizational goals. They provide targeted focus, enhance efficiency and cost reduction, improve quality, enable continuous improvement, and facilitate performance measurement.

At the activity level, objectives bring clarity, prioritize tasks, mark milestones, ensure accountability and measurement, and promote continuous improvement.

At the development and innovation level, objectives foster creativity, exploration, and the generation of novel solutions within data science and research projects. This level encourages pushing boundaries, embracing risk-taking, and promoting out-of-the-box thinking.

Firm Performance Evaluation, the proposition discusses the evaluation of firm performance at different levels: business level, process level, activity level, and development & innovation level.

At the business level, key considerations include defining key performance indicators (KPIs) that align with strategic objectives, comparing performance against industry benchmarks and competitors, and gathering qualitative data through customer surveys and stakeholder feedback.

At the process level, evaluation involves defining process-related KPIs, collecting and analyzing data on process performance, and regularly tracking progress to identify areas for improvement and drive continuous enhancement.

At the activity level, the evaluation focuses on establishing clear objectives for each activity, tracking milestone achievement, and defining task-specific KPIs to evaluate the effectiveness and efficiency of individual tasks.

At the development & innovation level, evaluation aims to measure the organization's ability to foster creativity and innovation. It includes defining metrics for innovation outcomes, assessing the organization's learning and adaptation capabilities, and gathering stakeholder feedback.

By applying these evaluation elements and methods at each level, organizations can comprehensively assess firm performance, align strategies, improve operational workflows, enhance task execution, foster innovation, and drive continuous improvement.

Capability development: The proposition discusses the importance of **capability development** in organizations to leverage data and analytics for strategic decision-making and sustainable growth. It is divided into four levels: business, process, activity, and development & innovation.

At the business level, capability development involves aligning talent strategy, data strategy, technology infrastructure, and stakeholder engagement to drive data-driven decision-making and foster innovation. It emphasizes the need for technical and subject knowledge talent, a comprehensive data strategy, robust technology infrastructure, and engaging senior management and stakeholders.

At the process level, capability development focuses on integrating data science and research methods into existing processes to optimize efficiency, enhance decision-making, and drive continuous improvement. It highlights process optimization, continuous improvement, and adopting agile methodologies to utilize data science and research capabilities effectively.

At the activity level, capability development aims to enhance the skills, workflows, and documentation practices related to data science and research activities. It emphasizes skill enhancement, standardized workflows, and documentation and knowledge sharing to maximize the effectiveness and efficiency of individual activities within data science projects.

At the development & innovation level, capability development fosters a culture of innovation and enables organizations to leverage data science and research for transformative advancements. It involves encouraging innovation initiatives, collaborating with external partners, and managing intellectual property to drive breakthrough ideas, engage in joint research projects, and capitalize on innovative solutions.

Capability development at these four levels allows organizations to harness the power of data science, research, and business acumen, make informed decisions, drive innovation, and achieve sustainable growth.

Strategic alignment in data science and research projects provides insights into the barriers and enablers at different levels (Business Level, Process Level, Activity Level, and Development & Innovation Level).

At the Business Level, strategic alignment involves aligning data science and research projects with the organization's strategic objectives. Barriers include a lack of top-level management support, resistance to change, and limited stakeholder engagement.

Enablers include strong leadership commitment and clear communication of project benefits.

At the Process Level, strategic alignment integrates data science and research methodologies into existing processes. Barriers include resistance to change and lack of standardization. Enablers include leadership support, clear guidelines, and collaboration between process owners and data science teams.

Strategic alignment is crucial for executing specific data science and research activities at the Activity Level. Barriers include limited availability of skilled professionals and lack of resources. Enablers include talent acquisition programs, investment in data analytics tools, and proper resource allocation.

Strategic alignment at the Development & Innovation Level fosters development and innovation in data science and research projects. Barriers include a non-supportive organizational culture and insufficient collaboration. Enablers include an innovation-friendly culture, collaboration with external partners, and intellectual property management.

By addressing these barriers and leveraging the enablers, organizations can achieve strategic alignment, maximize the value of data-driven decision-making, and drive successful outcomes in data science and research projects.

Table 3. Summary

Proposition Category	**Subcategory**	**Summary**
Objective	Business Level	The objective setting at the business level is crucial for data science and research projects. It involves defining clear and measurable objectives that align with the organization's strategic goals. These

		objectives provide direction and purpose for the organization's data science initiatives, ensuring that efforts are focused on achieving tangible outcomes that contribute to overall business success.
	Process Level	The objective set at the process level focuses on defining specific goals and targets for the organization's data science and research processes. It involves breaking down high-level objectives into actionable targets that can be measured and monitored. By establishing clear process-level objectives, organizations can track the performance and effectiveness of their data science initiatives, identify areas for improvement, and make data-driven decisions to optimize their processes.
	Activity Level	The objective set at the activity level involves setting goals and targets for individual data science and research activities. It includes defining specific deliverables, milestones, and performance indicators for each activity, ensuring that they are aligned with the organization's broader objectives. By setting clear activity-level objectives, organizations can enhance accountability, track

		progress, and ensure that individual tasks contribute to the overall success of data science projects.
	Development & Innovation Level	The objective set at the development and innovation level aims to establish goals and targets for fostering a culture of continuous learning, development, and innovation within the organization. This includes setting objectives related to skill development, knowledge sharing, and innovation initiatives. By defining clear development and innovation objectives, organizations can promote a growth mindset, encourage experimentation, and drive continuous improvement in their data science and research practices.
Firm Performance	Business Level	Performance evaluation at the business level involves assessing the impact of data science and research projects on firm performance. This includes evaluating financial indicators, such as revenue growth and cost savings, as well as non-financial indicators like customer satisfaction, operational efficiency, and innovation. By measuring the overall impact of data science initiatives on firm performance, organizations can gauge the

		effectiveness of their investments and make informed decisions to optimize business outcomes.
	Process Level	Performance evaluation at the process level focuses on measuring the organization's effectiveness and efficiency of data science and research processes. It involves assessing key performance indicators (KPIs) related to process efficiency, quality, and cycle time. By evaluating process-level performance, organizations can identify bottlenecks, streamline workflows, and continuously improve their data science and research processes to drive better outcomes and resource utilization.
	Activity Level	Performance evaluation at the activity level involves assessing the performance of individual data science and research activities. This includes evaluating factors such as accuracy, timeliness, and resource utilization for each activity. By measuring activity-level performance, organizations can identify strengths and weaknesses, allocate resources effectively, and ensure that individual tasks contribute to the overall success of data science projects.

	Development & Innovation Level	Performance evaluation at the development and innovation level aims to measure the impact of development and innovation initiatives on the overall performance of data science and research projects. It includes evaluating innovation initiatives' success, new technologies' adoption, and learning outcomes. By assessing development and innovation performance, organizations can foster a culture of continuous improvement, encourage experimentation, and drive innovation in their data science practices.
Capability Development	Business Level	Capability development at the business level focuses on enhancing the skills and knowledge of individuals involved in data science and research projects. This includes identifying competency gaps, defining learning objectives, and implementing training programs. By investing in developing data science capabilities at the business level, organizations can equip their workforce with the necessary skills to drive innovation, make informed decisions, and achieve better outcomes in their data science initiatives.

	Process Level	Capability development at the process level involves improving the organization's data science and research processes. It includes adopting best practices, implementing standardization, and optimizing workflows. By enhancing process-level capabilities, organizations can improve efficiency, quality, and consistency in their data science practices, leading to more reliable and impactful outcomes.
	Activity Level	Capability development at the activity level focuses on enhancing the execution of data science and research activities. It includes providing training, allocating resources effectively, and promoting collaboration. By strengthening activity-level capabilities, organizations can ensure that individual tasks are performed effectively, leverage available resources efficiently, and promote a collaborative environment that enhances the overall effectiveness of data science projects.
	Development & Innovation Level	Capability development at the development and innovation level aims to foster a culture of continuous learning, development, and innovation within the organization. This includes promoting a

		growth mindset, encouraging knowledge sharing, and creating opportunities for experimentation and innovation. By nurturing a development and innovation-focused culture, organizations can continuously enhance their data science capabilities, drive innovation, and stay ahead in a rapidly evolving landscape.
Strategic Alignment	Business Level	Strategic alignment at the business level involves aligning data science and research projects with the organization's overall strategic objectives. It requires overcoming barriers such as lack of management support, resistance to change and ensuring that data science initiatives are integrated into the organization's strategic planning processes. By achieving strategic alignment, organizations can ensure that their data science efforts are purposeful and impactful and contribute directly to achieving their broader strategic goals.
	Process Level	Strategic alignment at the process level focuses on integrating data science and research methodologies into existing processes within the organization. It requires addressing barriers such as resistance to change, lack of

		standardization and ensuring that data science practices are embedded seamlessly into the organization's operational workflows. By achieving process-level alignment, organizations can leverage data science to optimize their processes, make data-driven decisions, and drive continuous improvement.
	Activity Level	Strategic alignment at the activity level involves efficiently executing data science and research activities within the organization. It requires addressing challenges such as the limited availability of skilled professionals, insufficient resource allocation, and ensuring that individual tasks align with the broader strategic objectives. By achieving activity-level alignment, organizations can ensure that their data science activities contribute directly to achieving strategic goals, maximizing their overall impact and value.
	Development & Innovation Level	Strategic alignment at the development and innovation level aims to foster an organizational culture that encourages development and innovation in data science and research projects. It includes creating an environment that supports

		experimentation, rewards innovation, and provides opportunities for learning and growth. By achieving alignment at the development and innovation level, organizations can cultivate a culture of continuous improvement, drive innovation, and leverage data science to its full potential.

4.7 Conclusion

In conclusion, developing an enterprise ecosystem for data science and research projects is essential for organizations to harness the full potential of their data and drive innovation. Our analysis and discussion explored several key propositions: objective setting, firm performance evaluation, capability development, and strategic alignment. These propositions provide a comprehensive framework for organizations to establish a strong foundation and maximize the impact of their data science initiatives.

Objective setting is critical in providing direction and purpose to data science projects. By defining clear objectives at the business, process, activity, and development levels, organizations can ensure that their efforts are aligned with strategic goals and deliver tangible outcomes. It enables organizations to stay focused, measure progress, and make data-driven decisions to optimize their processes and achieve desired results.

Firm performance evaluation allows organizations to assess the impact of data science projects on overall business performance. By measuring financial and non-financial indicators, organizations can gauge their investments' effectiveness, identify improvement areas, and make informed decisions to optimize business outcomes. Performance evaluation at the process, activity, and development levels further enables organizations to identify bottlenecks, improve efficiency, and foster a culture of continuous improvement and innovation.

Capability development is crucial for building a skilled workforce and optimizing data science practices. By investing in capability development at the business, process, activity, and development levels, organizations can enhance the skills and knowledge of their employees, improve processes, and foster a culture of continuous learning and innovation. It enables organizations to stay competitive, adapt to changing technologies, and drive impactful outcomes through data science initiatives.

Strategic alignment ensures that data science projects are integrated into the organization's overall strategic objectives. Organizations can ensure that their data science efforts are purposeful and integrated and contribute directly to achieving strategic goals by achieving alignment at the business, process, activity, and development levels. Strategic alignment addresses barriers such as resistance to change, lack of standardization, and limited resource allocation, enabling organizations to leverage data science to optimize processes, make data-driven decisions, and drive continuous improvement.

In conclusion, developing an enterprise ecosystem for data science and research projects requires a holistic approach, encompassing objective setting, firm performance evaluation, capability development, and strategic alignment. By adopting these propositions, organizations can establish a strong foundation for their data science initiatives, optimize their processes, drive innovation, and achieve better outcomes. Embracing data science and research within the enterprise ecosystem is critical to unlocking the full potential of data and gaining a competitive edge in today's data-driven business landscape.

CHAPTER V:

DISCUSSION

5.1 Introduction

The chapter interprets and analyzes the results in light of the research objectives and existing literature. It provides an in-depth analysis of the findings, explores their implications, and relates them to the research questions.

It presents a comprehensive discussion of the findings and analysis obtained from the research conducted on enterprise ecosystem development for data science and research projects. The discussion aims to provide a deeper understanding of the key propositions identified in the previous chapters: objective setting, firm performance evaluation, capability development, and strategic alignment. By delving into these propositions, their subcategories, and their implications, this chapter sheds light on the essential elements required for successful enterprise ecosystem development in the context of data science.

The chapter also discusses any limitations of the study and suggests areas for further research.

5.2 Objectives of the Research

This research aims to address the challenges enterprises face in developing and implementing data science and research projects within their ecosystems. The study aims to provide insights and recommendations for successfully establishing a strategic ecosystem that supports using data science and research project outcomes in enterprises.

The research objectives are as follows:

1. To understand the challenges and opportunities associated with ecosystem development for data science and research projects.

2. To identify the key factors influencing the success of ecosystem development in enterprises.
3. To develop a comprehensive framework for ecosystem development that integrates business objectives, technology, processes, and capabilities.
4. To provide recommendations for implementing the framework and overcoming barriers to ecosystem development in enterprises.

5.3 Filling the Gaps in Existing Literature

This research contributes to the existing literature on ecosystem development for data science and research projects in enterprises by addressing several gaps that have been identified:

1. Limited comprehensive frameworks: Previous studies have often focused on specific aspects of data science and research projects, such as methodologies or technical tools. However, a lack of comprehensive frameworks guides organizations in developing ecosystems to support these initiatives. Our research fills this gap by providing a holistic framework encompassing various ecosystem development dimensions.
2. Neglect of organizational factors: While technical aspects of data science projects have received significant attention, there is a gap in understanding the organizational factors that influence the success of ecosystem development. Our research highlights the importance of objective setting, performance evaluation, capability development, and strategic alignment in fostering a conducive ecosystem. We offer valuable insights that complement existing literature by emphasizing these organizational dimensions.
3. Cultural aspects: Existing literature often overlooks the cultural aspects that shape ecosystem development. Organizational culture is crucial in embracing data

science and research, promoting continuous learning, and fostering innovation. Our research highlights the significance of cultural factors and their impact on ecosystem development, thus filling a gap in understanding the organizational and cultural dimensions of data science initiatives.

5.4 Key Findings and Analysis

Through our analysis and discussion, we have explored several key propositions that stem from the research objectives and address the identified gaps in the literature. These propositions provide a comprehensive framework for organizations to establish a strong foundation and maximize the impact of their data science initiatives:

5.4.1 Objective Setting

Objective setting is a crucial component of enterprise ecosystem development for data science projects. Through the analysis of various literature and discussions, it has been established that objective setting should occur at multiple levels within the organization: business level, process level, activity level, and development & innovation level. This hierarchical approach ensures that objectives are aligned with the organization's strategic goals and enable focused efforts toward desired outcomes.

At the business level, objective setting involves defining overarching goals that align with the organization's mission and vision. These goals provide a strategic direction for data science initiatives and serve as a guiding framework for decision-making and resource allocation. Clear and well-defined objectives at this level allow organizations to align their data science efforts with their overall business strategy, ensuring that data-driven insights contribute to achieving broader organizational goals.

Moving down to the process level, objective setting focuses on establishing goals that pertain to specific data science processes and methodologies. It includes defining performance metrics, setting quality standards, and identifying key performance indicators

(KPIs) to measure the effectiveness and efficiency of data science processes. Organizations can identify bottlenecks, improve efficiency, and enhance operational effectiveness by setting process-level objectives, ultimately leading to better outcomes.

The objective set at the activity level involves defining goals for specific data science activities or tasks. It includes tasks such as data collection, cleaning, analysis, and model development. By setting clear objectives at this level, organizations can ensure that each activity contributes meaningfully to the overall objectives set at higher levels. These objectives provide a sense of purpose and direction to individual data science tasks, enabling teams to prioritize their efforts and allocate resources accordingly.

Lastly, at the development & innovation level, the objective setting focuses on fostering a culture of continuous improvement and innovation within the organization's data science ecosystem. This involves setting goals that promote exploring new techniques, adopting emerging technologies, and collaborating among data scientists. Objectives at this level encourage ongoing learning, experimentation, and integrating innovative practices into the organization's data science initiatives.

5.4.2 Firm Performance Evaluation

Firm performance evaluation is an essential aspect of measuring the impact and effectiveness of data science projects on overall business performance. By evaluating both financial and non-financial indicators, organizations can gain insights into the value generated by their data science and business initiatives.

At the process level, performance evaluation enables organizations to assess the efficiency and effectiveness of their data science and business processes. It includes evaluating factors such as turnaround time, accuracy of results, resource utilization, and adherence to quality standards. By conducting regular process-level performance evaluations, organizations can identify areas for improvement, optimize their processes, and ensure the delivery of high-quality outputs.

Similarly, at the activity level, performance evaluation provides a granular view of specific data science and business activities and their contribution to overall outcomes. It involves assessing the performance of individual tasks, such as data preprocessing, feature engineering, model training, validation, etc. By measuring the performance of these activities against predetermined objectives, organizations can identify strengths, weaknesses, and opportunities for enhancement.

Performance evaluation at the development & innovation level focuses on assessing the impact of innovative practices and integrating new technologies into the organization's data science ecosystem. It involves evaluating the success of experimentation, adopting new algorithms or tools, and generating novel insights. By assessing performance at this level, organizations can gauge the effectiveness of their efforts in driving innovation, fostering learning, and staying at the forefront of data science advancements.

Moreover, firm performance evaluation also includes the assessment of financial indicators, such as return on investment (ROI) and cost savings resulting from data science projects. By quantifying the financial impact of data-driven initiatives, organizations can justify investments, prioritize projects, and align their data science efforts with overall business objectives.

5.4.3 Capability Development

Capability development plays a crucial role in enabling organizations to build the necessary skills, expertise, and infrastructure required for successful enterprise ecosystem development in data science and research. By investing in capability development, organizations can enhance their data science capabilities, create a competitive advantage, and drive innovation.

At the business level, capability development involves establishing a strategic roadmap for building data science capabilities within the organization. It includes identifying the required skill sets, recruiting and training talent, and establishing partnerships or collaborations with external entities, such as universities or research institutions. By strategically developing capabilities at this level, organizations can ensure a strong foundation for their data science initiatives and align their workforce with the demands of the evolving data landscape.

Capability development at the process level focuses on enhancing the efficiency and effectiveness of data science processes. It includes investing in tools, technologies, and infrastructure that support data collection, storage, analysis, and visualization. By equipping data scientists with the necessary tools and resources, organizations can streamline their processes, reduce bottlenecks, and improve their data science projects' overall productivity and quality.

At the activity level, capability development involves providing training and development opportunities for data science professionals. It includes offering specialized training programs, organizing workshops or seminars, and encouraging knowledge sharing and collaboration among data scientists. By investing in the continuous development of individual skills, organizations can nurture a culture of learning, enhance the expertise of their workforce, and promote innovation within their data science ecosystem.

Capability development at the development & innovation level focuses on fostering a culture of experimentation, exploration, and continuous improvement. This includes creating platforms or environments that encourage data scientists to explore new techniques, algorithms, and technologies. By providing the necessary support and resources for experimentation and innovation, organizations can drive breakthroughs, facilitate the development of novel solutions, and stay ahead in the rapidly evolving field of data science.

5.4.4 Strategic Alignment

Strategic alignment is a critical factor for the success of enterprise ecosystem development in data science. By aligning data science initiatives with the organization's overall strategy, goals, and values, organizations can ensure that data-driven insights contribute directly to business outcomes and competitive advantage.

At the business level, strategic alignment involves aligning data science objectives with the organization's overall strategic goals and objectives. This includes identifying areas where data science can create value, support decision-making, and drive innovation. By aligning data science initiatives with strategic priorities, organizations can allocate resources effectively, prioritize projects, and ensure that data science efforts are directly linked to achieving broader organizational goals.

Strategic alignment at the process level focuses on integrating data science into existing business processes and workflows. This includes identifying touchpoints where data science can add value, streamlining data collection and analysis processes, and integrating data-driven insights into decision-making frameworks. By aligning data science with existing processes, organizations can leverage their data assets effectively, drive efficiency, and enable data-driven decision-making at all levels of the organization.

At the activity level, strategic alignment involves aligning specific data science activities with business objectives and priorities. This includes ensuring that data collection, analysis, and modeling activities are designed to address specific business challenges or opportunities. By aligning activities with business needs, organizations can focus their data science efforts on areas that offer the highest potential for impact and value creation.

Strategic alignment at the development & innovation level emphasizes the alignment of innovative practices, emerging technologies, and data science experimentation with the organization's strategic goals. This includes exploring new avenues for data-driven innovation, identifying emerging trends or technologies, and aligning data science research and development activities with the organization's long-term vision. By strategically aligning development and innovation efforts, organizations can ensure they stay at the forefront of data science advancements and maximize the value derived from their ecosystem.

5.5 Limitations and Future Research

While our research provides valuable insights, it is essential to acknowledge its limitations:

- The study focused on a specific industry, which may limit the generalizability of the findings. Future research should explore different industries and contexts to ensure broader applicability.
- The reliance on self-reported data introduces potential bias and subjectivity. Future research could employ objective measures and quantitative approaches for validation and expansion.
- The limited timeframe of the study may have impacted the depth of analysis and the ability to capture long-term outcomes. Longitudinal studies would provide a more comprehensive understanding of the impact of ecosystem development on organizations over time.
- Cross-cultural aspects of ecosystem development require further investigation. Comparative studies across different cultural contexts would enhance our understanding and provide valuable insights for organizations operating in diverse environments.

5.6 Conclusion

In conclusion, our research fulfills the gaps in the existing literature by providing insights into the challenges, success factors, and a comprehensive framework for ecosystem development in enterprises. The findings contribute to a holistic understanding of ecosystem development for data science and research projects, emphasizing the importance of organizational factors and cultural aspects. By bridging these gaps, our research offers practical implications for both academia and practitioners.

Future research should build upon these findings, addressing the identified limitations and further enhancing our understanding of ecosystem development and its impact on organizations. By embracing data science and research within the enterprise ecosystem, organizations can unlock the full potential of data, gain a competitive edge in today's data-driven business landscape, and drive innovation and success.

CHAPTER VI:

SUMMARY, IMPLICATIONS, AND RECOMMENDATIONS

6.1 Summary

This chapter comprehensively summarizes the research conducted on enterprise ecosystem development for data science and research projects. The research aimed to explore the various aspects of building an effective enterprise ecosystem to leverage the potential of data science and research for organizational success. The findings were organized into four main categories: Objectives, Firm Performance Evaluation, Capability Development, and Strategic Alignment. Each category consists of subcategories that further elucidate the research outcomes.

6.1.1 Objectives

Under the Objectives category, the research focused on identifying the key objectives organizations should consider when developing an enterprise ecosystem for data science and research projects. The subcategories included:

- Business-Level Objectives: This subcategory examined how aligning data science initiatives with the organization's overall strategic objectives can drive performance improvement and innovation.
- Process-Level Objectives: The research explored the importance of streamlining data science processes and establishing efficient workflows to enhance project delivery and outcomes.
- Activity-Level Objectives: This subcategory delved into the specific activities and tasks that contribute to the successful execution of data science projects and how they can be optimized for better results.

- Development & Innovation-Level Objectives: The research highlighted the significance of fostering a continuous development and innovation culture to drive long-term success in data science initiatives.

6.1.2 Firm Performance Evaluation

The Firm Performance Evaluation category focused on evaluating the impact of data science and research projects on organizational performance. The subcategories included:

- Financial Performance Evaluation: This subcategory explored how data-driven insights and predictive analytics can enhance financial performance measurement and forecasting accuracy.
- Operational Performance Evaluation: The research examined the role of data science in improving operational efficiency, quality control, and supply chain management.
- Customer Performance Evaluation: This subcategory investigated the use of data science techniques to analyze customer behavior, preferences, and sentiment, enabling organizations to enhance customer satisfaction and loyalty.
- Employee Performance Evaluation: The research explored how data science can be leveraged to evaluate employee performance, identify skill gaps, and personalize learning and development initiatives.

6.1.3 Capability Development

The Capability Development category focused on building and nurturing the necessary capabilities within the organization to support data science and research projects. The subcategories included:

- Talent Acquisition and Management: This subcategory examined strategies for attracting top data science talent, building diverse teams, and implementing effective talent management practices.
- Skill Development and Training: The research highlighted the importance of continuous learning and development opportunities to enhance the skills and competencies of data scientists and other stakeholders involved in data science projects.
- Infrastructure and Technology: This subcategory explored the infrastructure requirements and technology investments necessary to support data science initiatives, including data storage, processing capabilities, and advanced analytics tools.

6.1.4 Strategic Alignment

The Strategic Alignment category focused on aligning data science and research projects with the organization's overall strategic direction. The subcategories included:

- Organizational Alignment: The research examined how organizations can align their data science initiatives with their vision, mission, and strategic goals, ensuring that data science projects contribute to the organization's overall success.
- Stakeholder Alignment: This subcategory explored the importance of engaging and aligning stakeholders from various departments and levels of the organization to ensure their support and commitment to data science initiatives.
- Collaborative Alignment: The research highlighted the significance of fostering collaboration and partnerships with external stakeholders, such as academic institutions, industry experts, and research organizations, to leverage their expertise and resources.

6.2 Implications

The implications section discusses the practical implications of the research findings for organizations aiming to enhance their enterprise ecosystem for data science and research projects. These implications consider the specific subcategories within each main category and provide actionable insights for organizations to consider.

6.1.5 Objectives:

- Business Level: Clearly defining the strategic objectives for data science and research projects and aligning them with the overall organizational goals is crucial. Organizations should articulate specific business outcomes they aim to achieve through these projects, such as improving customer satisfaction, enhancing operational efficiency, or driving innovation.
- Process Level: Organizations need to establish streamlined processes and workflows to support the execution of data science and research projects. It involves defining clear roles and responsibilities, establishing communication channels, and integrating data science activities into existing organizational processes.
- Activity Level: Identifying and prioritizing the key activities required for successful data science and research projects is essential. It includes data collection, preprocessing, modeling, analysis, and reporting. Organizations should allocate resources and manage these activities effectively to ensure timely and high-quality project deliverables.
- Development & Innovation Level: Encouraging a culture of continuous learning and innovation is critical. Organizations should invest in professional development opportunities for data scientists and researchers, promote collaboration and knowledge sharing, and create an environment that fosters experimentation and creativity.

6.1.6 Firm Performance Evaluation:

- Business Level: Organizations need to establish performance metrics that align with their strategic objectives and measure the overall impact of data science and research projects on firm performance. It may include metrics such as revenue growth, cost savings, customer retention, or market share.

- Process Level: Implementing robust performance evaluation mechanisms at the process level allows organizations to assess the efficiency and effectiveness of data science and research processes. Key performance indicators (KPIs) related to data quality, project timelines, resource utilization, and stakeholder satisfaction can be used to evaluate process performance.
- Activity Level: Monitoring and evaluating individual activities within data science and research projects is important to identify bottlenecks, improve efficiency, and ensure adherence to best practices. Performance metrics at the activity level may include the accuracy of predictive models, data completeness, data security, and adherence to ethical guidelines.
- Development & Innovation Level: Measuring the impact of development and innovation efforts is critical. Organizations should assess the effectiveness of their initiatives in terms of new product/service development, intellectual property creation, adoption of advanced technologies, and collaboration with external partners.

6.1.7 Capability Development:

- Business Level: Developing a strategic plan for building data science and research capabilities is essential. This involves assessing current capabilities, identifying skill gaps, and designing targeted training and recruitment strategies to attract and retain top talent.
- Process Level: Organizations should establish processes for identifying and acquiring the necessary data science tools, technologies, and infrastructure to support capability development. This includes

evaluating vendor solutions, implementing data management systems, and ensuring organizational processes are compatible.

- Activity Level: Providing ongoing training and development opportunities for data scientists and researchers at the activity level is crucial. This can include workshops, seminars, certifications, and collaboration with academic institutions or industry experts to enhance technical skills, domain knowledge, and research methodologies.
- Development & Innovation Level: Organizations should foster an environment that encourages development and innovation. This can be achieved by establishing dedicated research and development teams, incentivizing idea generation, and providing resources and support for experimenting with new approaches and technologies.

6.1.8 Strategic Alignment:

- Business Level: Aligning data science and research projects with the broader organizational strategy is essential. Organizations should ensure that these projects directly contribute to achieving strategic goals and that resource allocation and decision-making processes reflect this alignment.
- Process Level: Integrating data science activities into existing organizational processes and decision-making frameworks is crucial for strategic alignment. This involves incorporating data-driven insights into strategic planning, budgeting, resource allocation, and performance evaluation processes.
- Activity Level: Ensuring alignment at the activity level requires close coordination between data scientists, researchers, and business units. Organizations should establish mechanisms for effective communication, collaboration, and knowledge sharing to facilitate the translation of insights and research findings into actionable business strategies.

- Development & Innovation Level: Organizations need to align development and innovation efforts with strategic objectives. This can be achieved by defining research priorities, allocating resources accordingly, and fostering collaboration between development teams and business units.

6.3 Recommendations

Based on the research findings and implications, the following recommendations are proposed for organizations aiming to enhance their enterprise ecosystem for data science and research projects:

1. Develop a clear and comprehensive data science strategy aligning with the organization's strategic goals and objectives.
2. Foster a culture of innovation and continuous development to drive long-term success in data science initiatives.
3. Invest in attracting and retaining top data science talent and provide ongoing skill development and training opportunities.
4. Establish efficient data science processes and workflows to streamline project delivery and improve outcomes.
5. Implement data-driven performance evaluation mechanisms to measure and monitor organizational performance across various dimensions.
6. Leverage advanced analytics tools and technologies to support data science initiatives and enhance data processing and analysis capabilities.
7. Engage and align stakeholders at all levels of the organization to ensure their support and commitment to data science projects.
8. To leverage their expertise and resources, Foster collaboration and partnerships with external stakeholders, such as academic institutions and research organizations.

These recommendations provide a roadmap for organizations to effectively develop their enterprise ecosystem and leverage data science and research for innovation, performance improvement, and sustainable growth.

6.4 Conclusion

This chapter presented a summary of the research findings on enterprise ecosystem development for data science and research projects. The implications and recommendations provided valuable insights and guidance for organizations seeking to enhance their data science capabilities and leverage the power of data-driven decision-making.

APPENDIX A

SURVEY COVER LETTER

Dear [Participant name],

I hope this message finds you well. I am writing to inquire if you would be interested in participating in my doctoral research by granting me an interview. My dissertation focuses on developing an "Ecosystem development framework for enterprise data science and research projects." With your experience in leading data science and research projects in [xyz company or developing projects], you would be an ideal participant for my study.

My research aims to investigate how data science leaders create ecosystems for data science and research projects in enterprises. I am seeking interview participants who meet at least two of the following criteria: 1) a minimum of 3 years of experience in data science or research areas, 2) the successful execution of at least one data science project, and 3) experience leading a team of data scientists or researchers.

I value your time and will ensure our interview is as efficient as possible. At your convenience, the research will require a few hours of your time over the next few months. The initial interview is expected to take between 50 and 120 minutes, followed by a shorter follow-up interview to obtain feedback on the model as it emerges.

I understand that confidentiality may be a concern. I want to assure you that I will not collect proprietary information during the interview, and all personal identifying information will be removed from the results.

Your participation in this study would be immensely valuable and contribute to a substantial area of the scholarly literature. Moreover, it would help establish an ecosystem development framework for data science and research projects that is relevant and valuable to industry leaders. I would be delighted to present my final results to you and/or your organization.

Please let me know if you are willing to participate. I am happy to discuss any questions or concerns you may have over a quick phone call before committing.

Thank you for considering my request.

Best regards, [Researcher's Name]

APPENDIX B

INFORMED CONSENT

Interview Consent Form

Research project title: ECOSYSTEM DEVELOPMENT FRAMEWORK FOR ENTERPRISE DATA SCIENCE & RESEARCH PROJECTS

Research investigator: Shitalkumar R. Sukhdeve

Research Participants name:

The interview will take 60 to 90 minutes. We don't anticipate that there are any risks associated with your participation, but you have the right to stop the interview or withdraw from the research at any time.

Thank you for agreeing to be interviewed as part of the above research project. Ethical procedures for academic research require that interviewees explicitly agree to be interviewed and how the information contained in their interview will be used. This consent form is necessary for us to ensure that you understand the purpose of your involvement and that you agree to the conditions of your participation. Would you, therefore, read the accompanying **information sheet** and then sign this form to certify that you approve the following:

- the interview will be recorded and a transcript will be produced
- you will be sent the transcript and allowed to correct any factual errors
- the transcript of the interview will be analyzed by **Shitalkumar R. Sukhdeve** as the research investigator
- access to the interview transcript will be limited to **Shitalkumar R. Sukhdeve** and academic colleagues and researchers with whom he might collaborate as part of the research process
- any summary interview content, or direct quotations from the interview, that are made available through academic publications or other academic outlets will be **anonymized** so that you cannot be identified, and care will be taken to ensure that other information in the interview that could identify yourself is not revealed
- the actual recording will be destroyed after the submission of the research.

- any variation of the conditions above will only occur with your further explicit approval

Or a quotation agreement could be incorporated into the interview agreement

Quotation Agreement

I also understand that my words may be quoted directly. With regards to being quoted, please initial next to any of the statements that you agree with:

	I wish to review the notes, transcripts, or other data collected during the research pertaining to my participation.
	I agree to be quoted directly.
	I agree to be quoted directly if my name is not published and a made-up name (pseudonym) is used.
	I agree that the researchers may publish documents that contain quotations by me.

All or part of the content of your interview may be used;

- In academic papers, policy papers, or news articles
- On our website and in other media that we may produce such as spoken presentations
- On other feedback events
- In an archive of the project as noted above

By signing this form I agree that;

1. I am voluntarily taking part in this project. I understand that I don't have to take part, and I can stop the interview at any time;
2. The transcribed interview or extracts from it may be used as described above;
3. I have read the Information sheet;
4. I don't expect to receive any benefit or payment for my participation;
5. I can request a copy of the transcript of my interview and may make edits I feel necessary to ensure the effectiveness of any agreement made about confidentiality;
6. I have been able to ask any questions I might have, and I understand that I am free to contact the researcher with any questions I may have in the future.

Printed Name

Participants Signature Date

Researchers Signature Date

Contact Information

This research has been reviewed and approved by the Edinburgh University Research Ethics Board. If you have any further questions or concerns about this study, please contact:

Name of
researcher:
Shitalkumar R.
Sukhdeve

Full address:
Gondia
(Maharashtra),
India
Tel: +91 9673394344
E-mail: srskumar.sukhdeve@gmail.com

You can also contact (Researchers name) supervisor:
- Name of researcher: Monika Singh
- E-mail: monika.singh@ssbm.ch

What if I have concerns about this research?

If you are worried about this research, or if you are concerned about how it is being conducted, you can contact SSBM by email at contact@ssbm.ch.

BIBLIOGRAPHY

Abdelbaki, W. et al., 2023. A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. *Journal of Innovation & Knowledge,* 8(1).

Akter, S. et al., 2016. *How to improve firm performance using big data analytics capability and business strategy.* [Online]
Available at: https://ro.uow.edu.au/cgi/viewcontent.cgi?article=2021&context=buspapers
[Accessed 12 05 2022].

Behl, A. et al., 2019. *A conceptual framework for the adoption of Big Data Analytics by ecommerce startups: A case-based approach.* [Online]
Available at: https://cronfa.swan.ac.uk/Record/cronfa51740
[Accessed 12 05 2022].

Borasi, P., 2021. *Data Warehousing Market Insights - 2028.* [Online]
Available at: https://www.alliedmarketresearch.com/data-warehousing-market
[Accessed 10 3 2022].

Borges , A., Laurindo , F., Spínola , M. & Gonçalves , R., 2021. The strategic use of artifcial intelligence in the digital era: systematic literature review and future research directions. *International Journal of Information Management,* 57(102225).

Businesswire, 2021. *Data Warehouse Costs Soar, ROI Still Not Realized.* [Online]
Available at: https://www.businesswire.com/news/home/20210812005242/en/Data-Warehouse-Costs-Soar-ROI-Still-Not-Realized
[Accessed 10 3 2022].

Ciampi, F., Marzi, G., Demi, S. & Faraoni, M., 2020. The big data-business strategy interconnection: a grand challenge for knowledge management. A review and future perspectives. *Journal of Knowledge Management,* Vol. 24(No. 5), p. 24(5):1157–1176.

Coghlan, D., 2019. *Doing action research in your own organization.* s.l.:Sage.

Creswell, . J. W. & Poth, C. N., 2018. *Qualitative inquiry and research design: Choosing among.* 4th ed. s.l.:Sage.

Creswell, J. W., 2007. *Qualitative inquiry and research design: Choosing among five.* 2nd ed. s.l.:Sage.

Dastgerdi, A. K. & Gandomani, T. J., 2021. On the Appropriate Methodologies for Data Science Projects. *International Conference on Information Technology (ICIT),* pp. 667-673.

Davenport, T. H., 2020. *An Emerging Consensus Among Chief Data Officers.* [Online]
Available at: https://www.forbes.com/sites/tomdavenport/2020/02/26/an-emerging-consensus-among-chief-data-officers/?sh=1bd8b2c11315
[Accessed 10 3 2022].

Davenport, T. H., 2020. *Are You Asking Too Much of Your Chief Data Officer?.* [Online]
Available at: https://hbr.org/2020/02/are-you-asking-too-much-of-your-chief-data-officer
[Accessed 10 3 2022].

Dick, B., 2005. *Grounded theory: A thumbnail sketch. Aral..* [Online]
Available at: http://www.aral.com.au/resources/grounded.html
[Accessed 11 3 2022].

Dubin, R., 1976. *Theory building in applied areas..* s.l.: In M. D. Dunnette .

Elia, G. P. G. S. G. &. P. G., 2019. A multi-dimension framework for value creation through big data. *Industrial Marketing Management..*

Elia, G., Polimeno, G., Solazzo, G. & Passiante, G., 2020. A multi-dimension framework for value creation through Big Data. *Industrial Marketing Management,* Volume 90, pp. 617-632.

Fast, E. M., 2021. *Digital commons@George Fox University.* [Online]
Available at: https://digitalcommons.georgefox.edu/dbadmin/52/
[Accessed 3 2022].

Gartner, 2019. *Why Only Half of CDOs Are Poised for Success.* [Online]
Available at: https://www.gartner.com/smarterwithgartner/half-of-cdos-succeed
[Accessed 10 3 2022].

Georgieva, I. L. C. T. T. e. a., 2022. From AI ethics principles to data science practice: a reflection and a gap analysis based on recent frameworks and practical experience. *AI Ethics 2,* p. 697–711.

Gerard, G., Lakhani, K. R. & Phanish, P., 2020. *What has changed? The Impact of Covid Pandemic on the Technology and Innovation Management Research Agenda.* [Online]
Available at: https://onlinelibrary.wiley.com/doi/10.1111/joms.12634
[Accessed 8 3 2022].

Gong, Y. & Janssen, M., 2021. Roles and Capabilities of Enterprise Architecture in Big Data Analytics Technology Adoption and Implementation. *Journal of theoretical and applied electronic commerce research,* vol.16( ).

Grossman, R. L., 2018. A framework for evaluating the analytic maturity of an organization. *International Journal of Information Management,* pp. 45-51.

Jägare, U., 2019. Data science strategy for dummies. In: *Data science strategy for dummies.* s.l.:.John Wiley and Sons.

Johannessen, J., 1996. Systemics applied to the study of organizational fields: developing a systemic research strategy. *Kybernetes,* 25(1), pp. pp. 33-50.

Johson, J. S., Friend, S. B. & Lee, H. S., 2017. *Big Data Facilitation, Utilization, and Monetization: Exploring the 3Vs in a New Product Development Process.* [Online]
Available at: https://doi.org/10.1111/jpim.12397
[Accessed 08 03 2022].

Kaggle, 2017. *The State of Data Science & Machine Learning.* [Online]
Available at: https://ailab-ua.github.io/courses/resources/the_state_of_data_science_machine_learning_-_kaggle_2017_survey.pdf
[Accessed 05 03 2022].

Kaggle, 2020. *2020 Kaggle Data Science & Machine Learning Survey.* [Online]
Available at: https://www.kaggle.com/paultimothymooney/2020-kaggle-data-science-machine-learning-survey
[Accessed 5 3 2022].

Karaveg, C., Thawesaengskulthai , N. & Chandrachai, A., 2014. Evaluation model for research and development commercialization capability. *Production & Manufacturing Research,* 2(1), p. 586–602.

L.Grossman, R., 2018. A framework for evaluating the analytic maturity of an organization. *International Journal of Information Management,* 38(1), pp. 45-51.

LaValle, S. et al., 2011. Big Data, Analytics and the Path From Insights to value. *MIT Sloan Management Review,* 52(2).

Mardiana, S., 2020. Modifying Research Onion for Information Systems Research. *Solid State Technology,* 63(4), pp. 1202-1210.

Mckinsey, 2017. *ARTIFICIAL INTELLIGENCE THE NEXT DIGITAL FRONTIER?.* [Online]
Available at:

https://www.mckinsey.com/~/media/mckinsey/industries/advanced%20electronics/our%20insights/how%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/mgi-artificial-intelligence-discussion-paper.ashx
[Accessed 10 3 2022].

Melnikovas, A., 2018. Towards an Explicit Research Methodology: Adapting Research Onion Model for Futures Studies. *Journal of Futures Studies,* 23(2), p. 29–44.

Mikalef, P., Boura,, M., Lekakos, G. & Krogstie, 2019. *Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment.* [Online]
Available at: https://doi.org/10.1111/1467-8551.12343
[Accessed 8 3 2022].

NewVantage, 2020. *NewVantage Partners Releases 2020 Big Data and AI Executive Survey.* [Online]
Available at:
https://www.businesswire.com/news/home/20200106005280/en/NewVantage-Partners-Releases-2020-Big-Data-and-AI-Executive-Survey
[Accessed 19 3 2022].

Patton, M. Q., 2002. *Qualitative research and evaluation methods. Thousand Oaks.* 4 ed. s.l.:Sage.

Rajesh, R. C., Biplab , B., Debasis, M. & Anandadeep, M., 2022. *A systematic literature review towards a conceptual framework for enablers and barriers of an enterprise data science strategy.* [Online]
Available at: https://link.springer.com/content/pdf/10.1007/s10257-022-00550-x.pdf
[Accessed 8 3 2022].

Rashmi Yogesh Pai, A. S. A. D. S. R. B. J. S. S. N. T. K. D. &. K. J. D., 2022. Integrating artificial intelligence for knowledge management systems – synergy among people and technology: a systematic review of the evidence. *Economic Research-Ekonomska Istraživanja,* pp. 7043-7065.

Reddy, R. B. B. M. D. e. a., 2022. A systematic literature review towards a conceptual framework for enablers and barriers of an enterprise data science strategy. *Information Systems and e-Business Management ,* p. 223–255.

Rene Abraham, J. S. J. v. B., 2019. Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management,* pp. 424-438.

Roberts, J., 2017. *4 reasons why most data science projects fail.* [Online]
Available at: https://www.ciodive.com/news/4-reasons-why-most-data-science-projects-fail/439637/
[Accessed 10 3 2022].

Saunders, M. L. P. &. T. A., 2016. *Research Methods for Business Students.* England: Pearson Education Limited.

Shad, M. K. et al., 2018. Integrating sustainability reporting into enterprise risk management and its relationship with business performance: A conceptual framework. *Journal of Cleaner Production,* Volume 208, pp. 415-425.

Strauss, A. & Corbin, J., 1994. *Grounded theory methodology: An overview..* Handbook of Qualitative Research ed. s.l.:In N. K. Denzin & Y. S. Lincoln (Eds.),.

Strauss, A. & Corbin, J., 1998. *Basics of qualitative research: Techniques and procedures for.* 2nd ed. s.l.:Sage.

Strauss, A. & Corbin, J., 2015. *Basics of qualitative research: Techniques and procedures for.* 4th ed. s.l.:Sage.

Sukhdeve, S. R., 2020. *Enterprise Data Science & Articial Intelligence with Big Data.* 1 ed. s.l.:KDP Amazon.

Wamba, S. F. et al., 2015. *How 'big data' can make big impact: Findings from a systematic review and a longitudinal case.* [Online]
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2577073
[Accessed 12 05 2022].

Wiener , M., Saunders , C. & Marabelli , M., 2020. Big-data business models: A critical literature review and multiperspective research framework. *J Inf Technol,* 35(1), p. 35(1):66–91.