Critical Success Factors Impacting Intelligent Process Automation—
A Data-first Machine Learning Approach

By

Baruri Venkata Kodanda Pani (111910005)

Dissertation
submitted in partial fulfilment of the requirements for the
Executive Fellow Programme in Management at the
Indian School of Business

Advisor: Abhishek Kathuria, Ph.D.

Committee Members: Prasanna Karhade, Ph.D., and Giri Kumar Tayi, Ph.D.

Abstract

Critical Success Factors Impacting Intelligent Process Automation—

A Data-first Machine Learning Approach

By: Baruri Venkata Kodanda Pani

Intelligent process automation (IPA) refers to a combination of emerging technologies that include artificial intelligence, machine learning (ML), and robotic process automation to automate and optimize business processes. IPA facilitates significant business process automation, which, in turn, enhances business value and leads to digital transformation for an organization. In this study, I seek to identify the success factors impacting IPA-led digital transformation.

I argue that is necessary to explain the complex patterns of factors that are critical for IPA success through a multilevel and multi-method investigation. Accordingly, I follow a three-stage research methodology, consisting of abduction through in-depth key informant interviews, decision-tree induction, and theory abduction to examine potential success factors that lie across four theoretical levels of analyses. The data set derived from a sample of 176 IPA projects from the financial services domain, implemented by a multibillion-dollar global IT service company, forms the basis for this data-first ML investigation.

I draw on several theoretical perspectives and qualitative interviews to identify predictors of IPA success from multiple levels, such as domain, business process, technology, and governance. Firstly, I utilize decision-tree induction to examine three dependent variables corresponding to different dimensions of IPA success—Full Time Equivalent Reduction, Process Efficiency Improvement, and Process Accuracy Improvement. I choose these three dimensions based on the literature review and elite informant interviews.

Secondly, I combine these three dependent variables into a formative construct, which I term *IPA Success Index* that holistically captures the extent of IPA success. Decision-tree induction is also served as the methodology to examine the predictors of the IPA Success Index. Apart from the findings of these examinations, this study generates domain-specific rules, offers theoretical insights, and develops generalizable theoretical propositions.

Thirdly, I investigate in-depth the relationships between the key success factor(s), identified in the first two studies, and IPA success using econometric analysis. This analysis helps validate the configurational causality obtained through the abduction–induction–abduction theory development framework with the potential outcome view of causality.

In this dissertation, I contribute sixteen rules, six insights and six propositions that unveil the critical success factors for IPA implementation success, which are validated by econometric results. The identified success factors, the IPA Success Index, and theoretical propositions from my dissertation contribute toward the growing literature on intelligent information systems within the larger stream of IT business value research. Managers and organizations across the globe shall benefit from these studies, thus allowing them to maximize the benefits of IPA-led digital transformation.

Acknowledgments

I would like to extend my deepest appreciation to my advisor, Prof. Abhishek Kathuria, for his unwavering encouragement, invaluable guidance, and infinite patience throughout my doctoral journey. His expertise, insightful feedback, and constructive criticism have immensely contributed to shaping my research and writing in innumerable ways.

I am grateful to my committee members, Prof. Prasanna Karhade, and Prof. Giri Tayi, for their precious guidance, continuous support, and critical review, which have been pivotal in advancing my research.

I am grateful to the esteemed Indian School of Business (ISB) for providing me with an opportunity to pursue the EFPM program, access to world-class professors and resources, and other support.

Additionally, I would like to acknowledge the invaluable contributions of my colleagues and friends at ISB, whose supportive and intellectually stimulating environment has been instrumental in my research progress. Their constructive feedback, intellectual discussions, and generosity have helped me refine my ideas and broaden my perspective.

Lastly, I want to express my deepest appreciation to my family, whose love, encouragement, and unwavering faith in me have sustained me through the numerous challenges and obstacles that I faced. Their unwavering support and trust in my abilities have been the cornerstone of my success, and I will always be grateful for their unwavering presence in my life.

Table of Contents

List of Figures

List of Tables

1 INTRODUCTION

Intelligent process automation (IPA) has a significant impact on digital transformation by accelerating and enhancing the digital transformation journey. IPA is a combination of technologies including artificial intelligence (AI), machine learning (ML), and robotic process automation (RPA) used for automating and optimizing business processes. This technology transforms traditional manual processes and thus enables businesses to automate repetitive and time-consuming tasks, freeing up employees to focus on higher value work.

1.1 Intelligent Process Automation and Digital Transformation

Some of the ways IPA impacts digital transformation are given as follows:

1) Increased efficiency and productivity: By automating repetitive and time-consuming tasks, IPA has increased efficiency and productivity in businesses. This enables businesses to focus on more important tasks that require human expertise and creativity, leading to better customer experiences and business outcomes.
2) Improved accuracy and quality: With the use of AI and ML, IPA has helped improve the accuracy and quality of business processes. This has reduced errors and improved the quality of the final data (output) used in decision-making.
3) Enhanced decision-making: IPA has helped businesses to make better and faster decisions by providing real-time insights and analytics. This has enabled businesses to identify patterns and trends and to make data-driven decisions and thus to better outcomes and higher levels of satisfaction.
4) Reduced costs: IPA has helped reduce costs by automating repetitive tasks that were previously performed manually. This has resulted in lower labor costs and reduced errors, which, in turn, have led to better cost management.

5) Improved scalability: IPA has made it easier for businesses to scale up their operations. By automating processes, businesses can handle more work without having to hire more employees, leading to better scalability and growth.

An example of IPA-led digital transformation could be seen in the finance department of large organizations. The finance department traditionally handles a high volume of data-processing tasks, which are time-consuming, repetitive, and prone to human errors. By implementing IPA, the finance department can automate many of such tasks and free up time for its employees to focus on higher value-added activities.

In this scenario, IPA could be used to automate tasks such as invoice processing, payment processing, and financial data entry. AI and ML algorithms could be used to analyze data and identify patterns, leading to more accurate financial forecasting and risk management. By implementing IPA, the finance department can reduce errors, improve efficiency, and reduce costs, while also improving overall financial performance.

Overall, IPA has been a game changer for digital transformation, enabling businesses to achieve their goals faster and more efficiently.

1.2 Growth of IPA

The intelligent process automation (IPA) market has been growing rapidly in recent years and is expected to continue to grow in the coming years. According to a report by Grand View Research, the global IPA market size was estimated at USD 10.3 billion in 2020 and is expected to grow at a compound annual growth rate (CAGR) of 13.4% from 2021 to 2028.

The market is driven by several factors including the need for businesses to improve efficiency, reduce costs, and enhance customer experience. The COVID-19 pandemic has also accelerated the adoption of IPA as businesses have had to rapidly adapt to remote work and digitize their operations.

In terms of technology, robotic process automation (RPA) is the largest segment of the IPA market, accounting for more than half of the market share. However, the use of artificial intelligence (AI) and machine learning (ML) in IPA is also growing rapidly and is expected to become increasingly important in the coming years.

In terms of region, North America is expected to be the largest market for IPA, followed by Europe and Asia Pacific. It is reported that the widespread presence of IT and telecom companies, as well as the increasing adoption of automation solutions, in healthcare and financial services sectors are driving the growth of the IPA market in North America.

In terms of industry, the banking, financial services, and insurance (BFSI) sector is the largest consumer of IPA solutions due to the high volume of repetitive and rule-based processes in the industry. However, other industries such as healthcare, manufacturing, and retail are also expected to see significant growth in the adoption of IPA solutions.

1.3 Recent Trends in Intelligent Process Automation

Intelligent process automation (IPA) is a rapidly evolving field, and several trends shape its development. Some of the key trends in IPA are as follows:

1) Increased adoption of AI and machine learning: AI and machine learning have become more prevalent in IPA and are used to create more sophisticated automation solutions. These technologies enable IPA to learn from data, adapt to changing conditions, and make more intelligent decisions.
2) Greater focus on hyperautomation: Hyperautomation is the use of multiple automation technologies, including RPA, AI, and machine learning, to automate complex processes

end-to-end. Hyperautomation enables businesses to achieve greater efficiency and agility, and better outcomes.

3) Growing use of low-code and no-code platforms: Low-code and no-code platforms are becoming more popular in IPA, enabling businesses to create automation solutions without needing extensive coding knowledge. These platforms enable easier and faster development and deployment of automation solutions.
4) Integration with other technologies: IPA is integrated with other technologies, such as the Internet of Things (IoT), blockchain, and cloud computing. This integration enables businesses to create more advanced automation solutions that are more scalable, flexible, and secure.
5) Emphasis on governance and compliance: As IPA has become more prevalent, there is a growing emphasis on governance and compliance. Businesses are required to ensure that their automation solutions are secure, compliant with regulations, and aligned with business objectives.
6) Focus on human-machine collaboration: IPA is not about replacing humans, but rather augmenting their capabilities. There is a growing focus on human-machine collaboration, where automation solutions work alongside humans to improve efficiency, productivity, and quality.

Overall, these trends shape the development of IPA and help businesses achieve greater efficiency, agility, and innovation.

1.4 How IPA Implementation Is Different from Other Software Development

Both software development and intelligent process automation (IPA) are important approaches to automating business processes and improving operational efficiency, but they differ in several keyways.

Software development typically involves building custom software solutions from scratch to meet specific business needs. This can be a time-consuming and resource-intensive process, but it can result in highly customized and tailored solutions that meet the exact requirements of a business. The development process typically involves a team of developers and other IT professionals who work together to design, develop, and test the software. For example, developing a new mobile app for a business may involve designing a user interface, coding the app functionality, testing it for bugs, and maintaining it over time.

By contrast, IPA involves using software tools and technologies to automate existing business processes, without the need for custom development. IPA solutions typically rely on pre-built software components, such as robotic process automation (RPA) bots, artificial intelligence (AI) algorithms, and machine learning models, that can be configured and integrated into existing systems to automate specific tasks or workflows. For example, an organization might use IPA to automate their customer service processes by implementing a chatbot that can handle routine inquiries and escalate more complex issues to human agents.

While software development can provide highly tailored and customized solutions, it can be time-consuming and expensive. IPA, on the other hand, can be faster and less expensive to implement, but may not be as customizable or flexible as custom software solutions.

Both software development and IPA can be used to automate business processes and improve operational efficiency. The choice between the two approaches will depend on a range of factors including the specific needs and requirements of the business, the complexity of the processes to be automated, and the resources available for development and implementation.

There have been many studies on the success factors of IT development projects. While the factors can vary depending on the specific study and context, some of the most identified success factors such as clear and well-defined project goals, effective project management, stakeholder engagement and support, skilled project team, adequate resources, agile project management approach, and effective quality control.

Intelligent process automation (IPA) is a relatively new technology, and there is limited research on the success factors specific to IPA projects. However, based on the available research and industry best practices, some key success factors of IPA projects are given in the following text.

1.5 Success and Failures of IPA Implementations

The intelligent process automation (IPA) market size is expected to grow significantly in the coming years. According to a report by Markets and Markets, the IPA market size was estimated at USD 10.0 billion in 2020 and is projected to reach USD 16.3 billion by 2025, with a compound annual growth rate (CAGR) of 10.2% during the forecast period. Considering the huge adoption of IPA raises the question of how successful these implementations are. The success percentage of IPA implementation can vary depending on several factors including the complexity of the processes being automated, the level of stakeholder engagement, and the quality of project planning and management.

However, industry studies have also shown that IPA implementation failure rates can be relatively high. For example, a 2021 study by Gartner found that by 2024, 50% of IPA implementations will fail to deliver sustained business value due to a variety of reasons, including a lack of expertise in process identification and automation, poor bot design and development, and inadequate change management and governance practices.

Another survey by Forrester Research in 2019 found that 30% of IPA projects stalled at the proof-of-concept stage, and only 16% of respondents reported achieving significant benefits from their IPA projects.

These figures indicate that IPA implementation can be challenging and that organizations must carefully plan and manage their IPA projects to avoid failure. It is essential to have a clear business case for automation, engage stakeholders early in the process, standardize processes, ensure data quality, plan for scalability and governance, and actively monitor bot performance to ensure that the automation is delivering the expected benefits. Additionally, organizations should focus on continuous improvement and actively address issues as they arise to ensure the success of their IPA implementations.

1.6 Purpose of this Dissertation Research

To improve the success percentage of IPA implementations, as discussed in section 1.5, examining the critical success factors of IPA implementation will be a huge advantage for organizations embarking on IPA journey, practitioners, and researchers in the field of information technology. In my dissertation research, I studied the live data of 176 IPA project implementations in banking and financial services.

This research extensively examined the success factors from the point of view of IPA. IPA is an important part of digital transformation (DT) for any organization, and it is prudent for the management to understand and strategize IPA implementations in such a way that they are successful in terms of selecting a right business process for automation, empowering business users with a right automation approach, identifying human interventions, etc. For example, designing completely attended IPA bots is suitable for simple repetitive tasks, while complex tasks may need a combination of bots and humans to run the business process effectively. There is a debate as to how the automation approach should be, i.e., top-down, or bottom-up.

The previously discussed factors pose a very important research area and have led to this research study on ***"Identify the success factors of Intelligent Process Automation"*** and answers the following two very important questions:

Question 1: What are the critical success factors that predict the success and failure of intelligent process automation implementation?

Question 2: What is the order of importance of the critical success factors that predict the success or failure of intelligent process automation?

These fundamental, yet complex, questions have strong theoretical and managerial implications, especially for firms that drive digital transformation through IPA. Although the research has examined several success factors for IT software development, there is limited empirical research on IPA. This necessitates a multi-level theoretical investigation as success of IPA is dependent on various constructs across multiple levels of analysis: IPA outcomes, how the IPA is governed, process-level constructs, technology considerations and complexity involved, and research approach.

Since IPA success factors encompass multiple levels, and each level often has (multiple) constructs, a multi-level approach is necessary to investigate the influence of the constructs and their emergent interactions across multiple levels (Paruchuri et al., 2018), while also requiring the integration of multiple theories as a single, homogeneous theory cannot be applied across levels (Hitt et al., 2007).

Furthermore, IPA success factors are not a single decision but encompass patterns of decision sequences. This decision journey, namely, the partial orderings of its constituent decision point, and decision forks are as critical as the final decision outcome itself.

This research has attempted to address this research gap by leveraging a multi-level, multi-theoretic approach that broadens notions of emergence in decision-making logics (Markus et al., 2002) by incorporating recent guidance on multi-level theorizing. A unique sample data set using an inductive data-driven analytics methodological approach was analyzed. Decision tree induction was used to identify patterns in the data and as a vital input in the abduction process, which generalizes the patterns to the most plausible explanation. In this study, the sample was 176 live IPA implementations in the banking and financial services sector across the globe executed by 200 global IT service providers. A three-stage research design of abduction-induction-abduction was applied, where the data derived from the live implementations were abducted and introduced into decision trees to derive the rules and abduct away from the rules to present insights and propositions.

By leveraging eleven predictors across governance, process, technology, and complexity-level perspectives, the impact of the predictors on the overall success or outcomes of intelligent process automation (IPA) was investigated. Thus, a key strength of this research design and study is that data are derived from real-time live implementations of 176 projects in the banking and financial services domain and it addresses heterogeneity as the data cover IPA implementations across the globe.

The question investigated in this study is important for IT managers and key executives to understand the factors that would impact the overall success of IPA implementations and hence design the overall program.

Decision trees can shine the light on the flow of the decision-making process and model the IPA success factors from the implementations studied in the banking and financial services along with their cumulative experiences.

In this study, decision trees were grown using C4.5 decision tree classification algorithm induction (Quinlan, 1986b, Quinlan, 1990) and were aggressively pruned to discover the underlying tacit structure of the data. Only few previous studies have utilized this methodology (Brézillon et al., 2002, Tessmer, 1994).

I then abduct away to discover theory, to articulate insights and generic propositions from the rules derived from decision tress for identifying factors impacting IPA success. In the abduction process, data were interpreted to discover combinations of features for which there is no appropriate explanation. This sequence of induction and abduction is appropriate as success factors of IPA implementation and combination of decision sequences it encompasses cannot be theorized ex-ante.

1.7 Key Findings

In this study, sixteen rules were derived for four IPA outcomes of success, and by abducting, six significant insights and six propositions were found. The predictors of successful IPA implementation were present at all theoretical levels, and there were dominant predictors for high and low success of IPA implementation. It was observed that there were only a few tacit combinations that result in successful IPA implementations. By observing the overall IPA implementation success, critical success factors that would impact the success of IPA implementations such as FTE reduction, process efficiency, and process accuracy were identified.

1.8 Contributions and Implications

In this study, predictors or critical success factors of IPA implementation were identified. Through the insights and propositions of this study, some dominant predictors that impact all the outcomes of successful IPA implantation were found. Furthermore, first- and second-level predictors of successful IPA implementation were determined. In some cases, specific combinations of predictors together impacting a specific outcome of IPA success were found, for example, process efficiency. In this study, decision trees help enlighten first principles, or "the essence of things," thereby representing a major contribution to theory. Progressive theory development through abduction reveals the intricate combinations of predictors and clarifies the influence of combinations of predictors on the outcome of interest.

This study offers a nuanced view into the decision-making process for IPA practitioners regarding the predictors or critical factors impacting both high and low IPA implementation success. It also offers organizing principles for their implementations by highlighting the most efficient path for increasing participation in their platforms and thus improving the probability of successful IPA implementation through rules, insights, and propositions. These practice implications also extend to other contexts of automation implementations such as low/code in other for healthcare, retail, and other sectors.

1.9 Conclusion

This chapter provides an introduction about the definition, growth, recent trends of intelligent process automation (IPA) and how it is different from other software development projects; discusses the success and failure of IPA implementation and what is the purpose of this dissertation; and finally discusses the reason as to why this topic and important research questions were selected.

The rest of this article is organized as follows: The second chapter will present a systematic literature review process, where the key factors pertaining to theory base, hypothesis, research methods, findings, and limitations of IPA lead digital transformation were extracted. Then, it will present the results of literature review of eight studies, followed by the antecedents of IPA led digital transformation and key thematic issues.

The third chapter will explain the research context and also the three-stage research methodology of abduction through hunches, induction through decision trees, and abduction thorough the examination of various outcome variables of IPA. Moreover, it will discuss how data are derived from the observed live implementations of IPA projects and elite informant interviews. Finally, the chapter will explain how twelve important predictors and three outcomes were derived from the observed data of 176 live IPA projects.

The fourth chapter will explain in detail theoretical levels (i.e., governance, process, technology, and complexity) and eleven predictors under the four theoretical levels in terms of their definitions and how they impact an IPA implementation.

The fifth chapter will explain the measures for this research study. First, it will detail the four outcomes of interest (i.e., key outcomes for a success or failure of an IPA project), namely, average handling time, FTE reduction, process efficiency, and process accuracy in terms of their definitions based on prior research and how the success is classified as high, medium, and low. Subsequently, it will explain the mechanism to classify, and code eleven predictors defined under the four theoretical levels impacting the outcomes of IPA success.

The sixth chapter will explain the induction mechanism using decision trees; describe how the computational experiments were performed; and describe the selection of a best representative tree after tree pruning based on three key heuristics, namely, parsimony, consistency, and prediction accuracy, which were examined independently based on the best representative tree; and present key findings and rules for each of the three outcomes: FTE reduction, process efficiency, and process accuracy. In addition, the chapter will present eleven rules representing critical success factors impacting the success or failure of the IPA implementation.

The seventh chapter will explain the difference between formative and reflective constructs and why the formative construct is justified, where a formative analysis was performed using principal component analysis (PCA) to arrive at one composite measure for IPA success from the three outcomes that were discussed in the sixth chapter. Following this, the IPA success composite measure was derived for each of the 176 IPA implementations by using the same process of decision tree induction and the same eleven predictors and by selecting the best representative tree using the three key heuristics to derive five rules for the critical success factors impacting the success or failure of an IPA implementation.

The eighth chapter will compare and contrast the rules derived from decision tree induction discussed in the sixth and seventh chapters to show the commonalities between the predictors impacting the outcomes of IPA implementation and observation and propositions for critical success factors of intelligent process automation (IPA).

The ninth chapter will explains the validation analysis through econometrics, where the direct effects of the top three predictors derived from decision tree induction and their combination effects were measured using pre-post analysis to statistically and empirically support the rules and propositions, i.e., impact of the important predictors before and after automation.

The tenth chapter will discuss the theoretical implications and contributions from the rules and propositions, managerial implications, strengths and limitations of the research, and the scope for future research and concluding remarks.

Figure 1 depicts the research roadmap for developing the thesis.

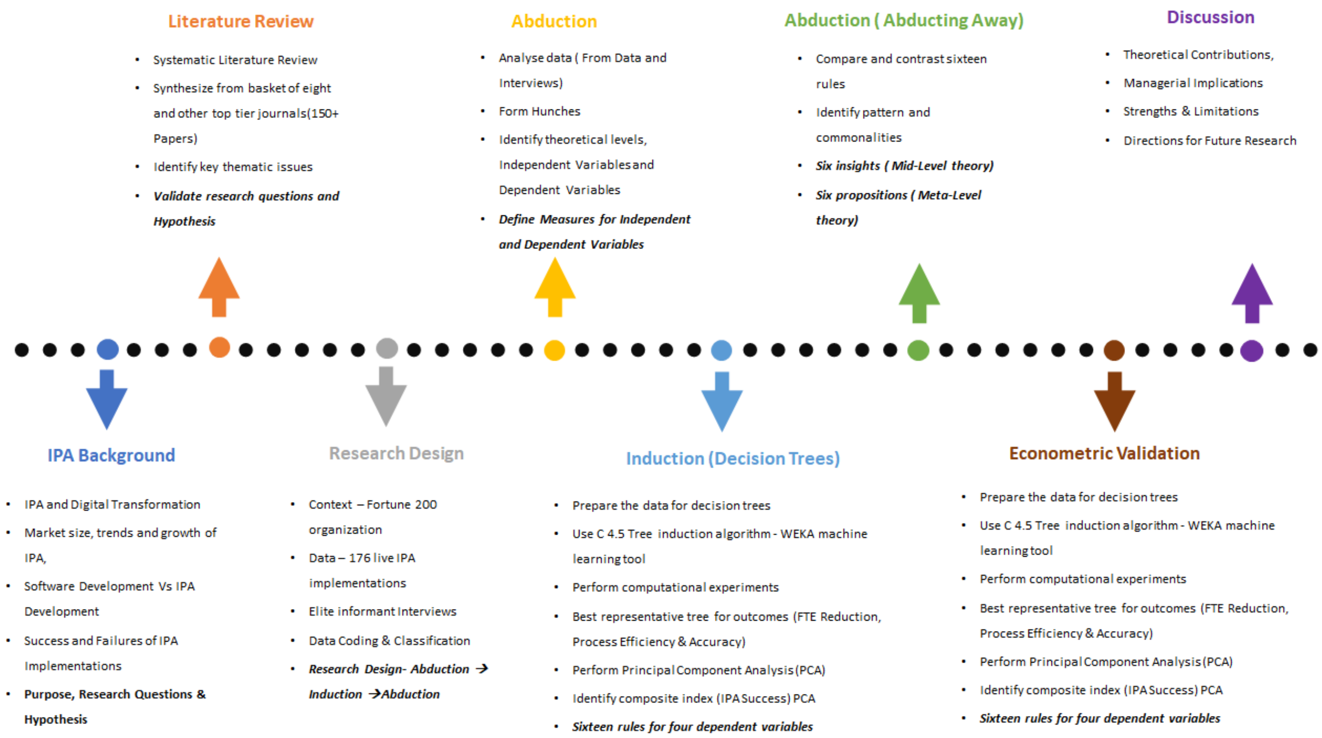

Figure 1: Research Roadmap

2 REVIEW OF LITERATURE

Multiple streams of theory inform this dissertation. In this study, the topic of interest is the theory of critical success factors that impact the intelligent process automation (IPA), a tool of digital transformation, and how it leads to digital transformation. This section presents a review of the extant literature by showcasing the major research themes in this area. First, the literature review process is explained, followed by the results of the literature review. Second, the conceptualization of the IPA-led digital transformation is described. Third, the antecedents and key thematic issues around IPA are outlined.

2.1 Systematic Literature Review Process

To meet the goals of this study, a comprehensive literature review has been performed, which involves various steps. The aim is to build theory, which improves the literature quality by synthesizing the concepts, rather than just reporting summary of various papers (Watson and Webster, 2020), which is relatively easy to understand. The synthesis of literature usually requires integrating concepts of interlinked topics for overall understanding of the intended paper or subject. The primary goal of literature review in the context of this study is to synthesize key themes, hypotheses, theory base, debates, and gaps in the extant literature (Templier and Pare, 2018, Vom Brocke et al., 2015).

In this review, I combine descriptive review and narrative review approaches as recommended by several literature review experts (Paré et al., 2015, Watson and Webster, 2020). A descriptive review approach is applied to a specific research area to unravel "any trends with respect to prior hypotheses, explainable patterns, theory base, research methodologies, or critical findings, while a narrative literature review approach combines and summarizes the extant literature to provide a holistic knowledge on a specific area of interest (Templier and Pare, 2018)

To conduct the literature review, AIS Senior Scholars' basket of eight journals were searched for articles published in the past twenty years. These journals are widely accepted as publishing high-quality research articles in the field of information systems. The basket of eight journals that were searched are *MIS Quarterly, Information Systems Research, Journal of Information Technology, Journal of Management Information Systems, Journal of Strategic Information Systems, Journal of the Association for Information Systems*, and *European Journal of information systems*, in addition to three other premier journals—"Organization Science", "Management Science," and "Strategic Management Journal," which have published significant knowledge about the focus area with respect to information systems. These journals constitute the premier journals of the reference fields of strategy and management.

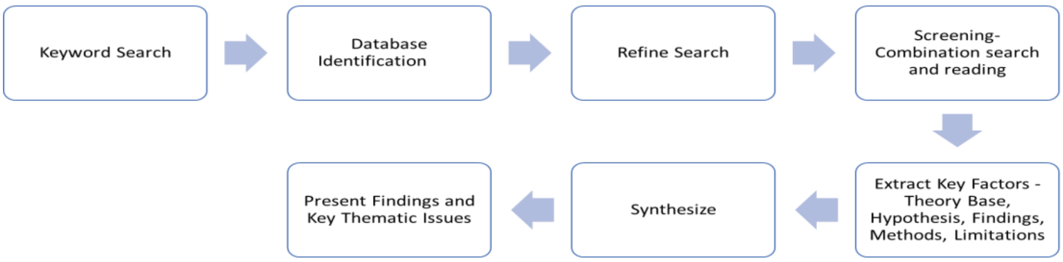

Figure 2: Literature Review Process

The elements considered for the literature search in this study were consistent with those previous scholars have used with respect to three broad focus areas: 1) digital transformation, 2) IT strategy in the context of *intelligent automation and robotic process automation*, and 3) where decision tree induction (used for theory induction and abduction) was used for theory induction through machine learning as a data-first approach. Accordingly, the following set of words were used to search for relevant research studies in the identified journals: [Digital Transformation, Intelligent Process Automation, Robotic Process Automation, Artificial Intelligence, Decision Trees etc.].

These terms were used both individually and in combination to get the relevant articles specific to the context of this study. After narrowing down the papers, the hypothesis, theory base, findings, and the limitations were generated for each element to gather information and present key thematic issues of this study. Figure 2 depicts the entire literature review process.

2.2 Framing literature Review in the Context of Intelligent Process Automation

Table 1 shows journal-wise break-up of articles reviewed in the context of intelligent process automation.

Table 1: Summary of results on intelligent process automation and digital transformation

Description	Count
Basket of eight (N=50)	
European Journal of Information Systems (EJIS)	7
Information Systems Research (ISR)	6
Journal of Strategic Information Systems (JSIS)	3
Journal of the AIS (JAIS)	6
Management Information Systems Quarterly (MISQ)	14
Journal of Management Information Systems (JMIS)	2
Information Systems Journal (ISJ)	2
Journal of Information Technology (JIT)	10
Other Significant Publications (N= 6)	
Management Science	2
Organization Science	2
Strategic Management Journal	2

2.3 Conceptualizing Digital Transformation

Digital transformation presents tremendous opportunities to organizations, information systems scholars, and practitioners. Although the extant literature has contributed to the understanding of digital transformation, there are still gaps in understanding of digital transformation initiatives. Organizations should understand that there is a difference between IT transformation and digital transformation because there is much investment that is being made into technology, business, and policy. There are many classical models of transformation, which would seem to undermine how digital transformation is different from IT-enabled transformation

(Henfridsson and Lyytinen, 2010). Many scholars attempted to explain the difference between IT and digital transformation; for example, (Vial, 2021) explained that the momentum for digital transformation would be larger, comprising "society and industry trends," whereas the momentum for IT transformation is more of managerial decisions. (Hartl and Hess, 2017) explained that the digital transformation is more complete and faster than IT transformation. Both these distinctions between digital transformation and IT transformation are related and are unclear as they cannot be proven empirically.

Digital transformation refers to *the integration of digital technology into all areas of an organization*, leading to *fundamental* changes to how the organization operates and delivers value to customers. It can enable organizations to become more efficient, agile, and customer centric. Thus, digital transformation impacts the core identity of firms. Research has examined how digital transformation impacts the organizational identity.

Organizational identity refers to the collective sense of purpose and shared values that define an organization and distinguish it from others. It is shaped by several factors including history, culture, brand, and mission. The process of digital transformation can have a significant impact on an organization's identity. For example, the adoption of new technologies may require an organization to change its culture and ways of working, which can, in turn, affect its sense of purpose and values. On the other hand, a strong organizational identity can help guide and inform the digital transformation process, ensuring that it aligns with the organization's core values and purpose.

It is important for organizations to consider the impact of digital transformation on their identity and to develop strategies for maintaining and reinforcing their identity throughout the process. This may involve engaging employees and stakeholders in the transformation process and ensuring that the organization's brand and values are consistently communicated and reinforced.

Organizational identity could be an effective and powerful way to conceptualize various transformations (Wessel et al., 2021). There are several examples that emphasize the significance of links between the organizational identity and value propositions. For example, Netflix transformed from a supplier of rental movies into a streaming platform, whereas Uber is a digital native that has transformed the whole car rental into a technology platform. There is a similarity between digital transformation and IT transformation in terms of technology effect on both organizational and environmental contexts. However, digital transformation is more about value-defining work in the organization, whereas IT transformation is all about value-supporting. Moreover, digital transformation creates a new identity for an organization, whereas IT transformation enhances the existing identity.

The extant literature emphasizes several impacts of digital transformation. (Baiyere et al., 2020) explained how digital transformation and business process management are often closely linked as digital technology can play a crucial role in supporting and enabling business process management initiatives. For example, digital tools and systems can help organizations automate and digitize their processes, reducing manual effort and increasing accuracy and speed.

Similarly, (Wimelius et al., 2021) explained a paradoxical perspective of technology renewal in digital transformation and highlighted the need for organizations to balance the need for innovation and change with the need for stability and continuity, and to navigate the challenges and risks associated with digital transformation.

(Tan et al., 2020) presented a very interesting view of the digital transformation of the K-pop industry, which has had a profound impact on its operations, marketing, and revenue generation, helping it to achieve significant growth and global recognition. It provides a compelling case study for organizations looking to transform their business ecosystems through the integration of digital technology.

Research on digital transformation also suggests several frameworks. (Gurbaxani and Dunkle, 2019a) recommended a framework for executives to assess their company's progress on six dimensions critical to successful digital transformation. Similarly, (Hess et al., 2016) proposed a conceptual framework for formulating a digital transformation strategy and key dimensions in terms of right questions to ask and provide managers with a comprehensive and structured approach to digital transformation.

There are also several challenges organizations encounter while implementing digital transformation. (Datta et al., 2020) clearly explained the challenges in terms of sociocultural disruption, digital literacy, and bureaucratic friction.

Overall, digital transformation refers to the integration of digital technology into all areas of an organization, fundamentally changing the way it operates and delivers value to customers. It is a strategic process that enables organizations to take advantage of the opportunities created by digital technology, such as increased efficiency, enhanced customer experience, and new business models.

Conceptualizing digital transformation involves considering the following key elements:

1. Technology: The use of digital technologies such as cloud computing, big data, artificial intelligence, and Internet of Things (IoT) to support and drive digital transformation.
2. Business processes: Rethinking and redesigning of business processes to optimize operations, reduce costs, and improve customer experience.
3. Culture and leadership: The alignment of organizational culture, leadership, and governance with the goals and objectives of digital transformation.
4. Data and analytics: The use of data and analytics to inform decision-making and optimize operations.
5. Customer engagement: The integration of digital technologies to enhance customer engagement and build stronger relationships with customers.

6. Innovation: The creation of new products, services, and business models enabled by digital technology.

Conceptualizing digital transformation involves considering the interplay between these elements and how they can be leveraged to drive strategic change and create value for the organization and its stakeholders. It also requires a holistic and integrated approach, considering the needs of all stakeholders and the impact of digital transformation on the wider business ecosystem.

2.4 Conceptualizing Intelligent Process Automation

Intelligent process automation (IPA) is a set of technologies and approaches that leverage artificial intelligence (AI) and machine learning (ML) to automate business processes. The goal of IPA is to automate repetitive, manual tasks and make work more efficient, accurate, and scalable.

Some common applications of IPA include chatbots for customer service, intelligent data capture and classification, predictive analytics, and robotic process automation (RPA). IPA can also be used to automate complex decision-making tasks and workflows, such as underwriting in the insurance industry or fraud detection in the financial services industry.

IPA combines the traditional benefits of process automation with the power of AI to create smarter, more flexible systems that can adapt to changing business needs. This can lead to improved accuracy, reduced cycle times, and increased efficiency (Carden et al., 2019) in areas such as customer service, finance, and human resources.

Denagama Vitharanage et al. (2020)) clearly illustrated "improvement in accuracy" as the most anticipated benefit of implementing IPA along with customer satisfaction.

It is important to note that while IPA can bring significant benefits, it is not a silver bullet solution and must be carefully integrated into an organization's existing processes and systems.

Effective implementation of IPA requires a well-defined strategy, clear business goals, and an understanding of the limitations and risks associated with the technology.

Asatiani and Penttinen (2016)) explained the challenges of IPA for Finnish financial firm OpusCapita in terms of how business processes are analyzed and assessed to arrive at a business case.

Overall, IPA represents a significant step forward in the evolution of process automation and has the potential to revolutionize the way work is done in many industries.

2.5 Intelligent Process Automation-Led Digital Transformation

IPA uses algorithms and software robots to mimic human-like decision-making and actions, thus enabling businesses to automate tasks that are repetitive, time-consuming, and prone to error.

Digital transformation, on the other hand, refers to the integration of digital technology into all areas of a business, resulting in fundamental changes to how the business operates and delivers value to customers. IPA plays a crucial role in digital transformation by enabling organizations to automate manual and time-consuming tasks, freeing up employees to focus on higher value activities, and delivering consistent, accurate, and scalable results.

For instance, organizations can use IPA to automate HR processes such as employee onboarding, payroll processing, and benefit administration, leading to improved efficiency and a better employee experience. IPA can also be used to automate customer service processes such as responding to customer queries and complaints, leading to faster and more efficient resolution times.

Table 2: IPA-led digital transformation – Exemplars

IPA success story 1: Credit limit extension	A bank received hundreds of credit limit extension requests daily. IPA and AI helped them achieve 100% accuracy and improve productivity by 91.67% in the limit extension of cash credit and overdraft facility. The process improvement had a high impact on the bank and end-customer credibility.
IPA success story 2: Bank account opening process automation	Due to the rapidly increasing number of customers, a bank wanted to automate their entire bank account opening process. RPA and OCR enable the bank to register a growth in the number of accounts opened per day from 3000 to 15000 with the same workforce and reduce processing time from 12 min to 3 min per case.
IPA success story 3: Auto-classification and auto-indexation of documents	For a US-based bank, which had acquired six banks, the number of documents in the document management system had increased manifold. RPA and AI/ML enabled the bank to auto-index and auto-classify ~35 million unstructured pages into 200+ categories.

In conclusion, IPA is a key technology that enables organizations to transform their operations and drive digital transformation by automating repetitive and manual tasks, freeing up employees to focus on higher value activities, and delivering consistent, accurate, and scalable results.

2.6 Antecedents of Digital Transformation and Intelligent Process Automation

The antecedents of digital transformation and intelligent process automation (IPA) can be traced back to several technological advancements and shifts in business practices. Some of the key antecedents as follows:

1. Increase in the Internet and cloud computing: The widespread availability of high-speed Internet and cloud computing has enabled organizations to access and store vast

amounts of data and information, which can be used to drive digital transformation initiatives.

2. Advances in artificial intelligence and machine learning: The advancements in AI and machine learning have made it possible to automate complex and time-consuming tasks, paving the way for the development of IPA.
3. Growth of mobile and Internet-connected devices: The proliferation of mobile devices and Internet of Things (IoT) has created a massive amount of data and made it possible to collect real-time information from various sources, providing organizations with new insights and opportunities for growth.
4. Need for agility and innovation: Organizations are under increasing pressure to respond to rapidly changing customer needs and market conditions. Automation and digital transformation enable organizations to be more agile and innovative, helping them to stay ahead of the competition.
5. Increasing focus on customer experience: With the increase in e-commerce and digital channels, customers have more options and higher expectations for the services and products they receive. Automation and digital transformation enable organizations to improve the customer experience and meet these increased expectations.

These antecedents have created the conditions for organizations to adopt digital transformation and IPA, leading to improved efficiency, cost savings, and enhanced customer experiences.

2.7 Key Thematic Issues

Although numerous studies in digital transformation-led intelligent process automation (IPA) offer significant insights into its conceptualization, they are mostly aimed at developing theory and not enough testing. Three themes emerge from the current review of the extant literature (see Appendix C for a list of indicative papers). First, prior research has focused heavily on guidelines, action principles, governance, and frameworks (Lacity et al., 2021, Kedziora and Penttinen, 2021, Lyytinen et al., 2021). For example, (Lacity et al., 2021) formalized an action principles approach for investigating and influencing the adoption of emerging information systems phenomena, particularly for new technologies such as IPA; it provides a six-step process of strategy, sourcing, program management, process selection, tool selection, and stakeholder buy-in as guidelines for IPA adoption. Kedziora and Penttinen (2021) explained governance models in Nordea Bank for the adoption of IPA and outlined several governance-related issues and decision points that must be addressed in connection with any deployment of intelligent process automation. Thus, studies have focused mainly on the general guidelines for RPA adoption through case study approaches and prior literature reviews but have not utilized actual data, which are the results of real-time IPA implementation. Therefore, in this study, 176 real-time IPA implementations that were executed in large banking and financial services domain across the globe.

The second issue to note is that intelligent process automation (IPA) is a recent phenomenon, and the success factors of intelligent automation have been used in multiple contexts, with considerable variation. (Oshri and Plugge, 2022) emphasized on process feasibility, service quality, and customer satisfaction. (Bygstad and Øvrelid, 2020) explained that deployment of lightweight IT in onsite configuration, loosely coupled with the infrastructure activities, allows for fast process innovation while leveraging the slow and nonlinear evolution of infrastructure. However, there is no clear articulation of success factors and benefit articulation. Therefore, the current study aims to quantitatively prove the success factors to be considered for IPA-led digital transformation by proving them empirically through machine learning and econometric research techniques and thereby establishing causality.

2.8 Conclusion

This chapter has detailed the literature review process and results, followed by key concepts of intelligent process automation (IPA) and digital transformation and how one leads to another, and key thematic issues identified from the literature review. The next chapter will describe the process of abduction derived from the 176 live implementations of IPA projects and elite informant interviews to arrive at the major predictors under the theoretical levels of governance, business process, technology, and complexity.

3 RESEARCH DESIGN

This study follows a three-stage research methodology consisting of abduction, induction, and abduction to generate multi-level theory for holistic understanding of phenomena through emergent combinations and sensemaking process (Figure 3).

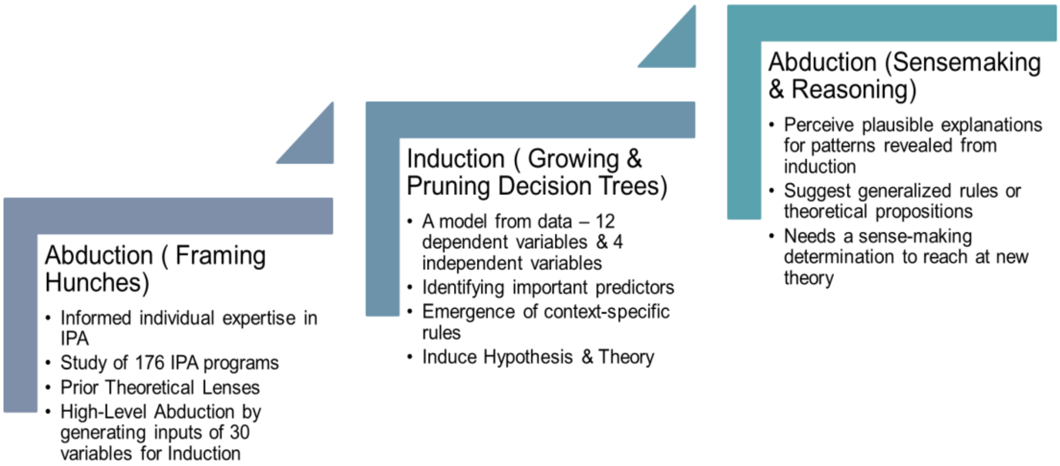

Figure 3: Research Methodology

In the first stage of this research methodology, as depicted in Figure 3, is abduction, hunches were generated and evaluated both at individual and collective levels (Sætre and Van de Ven, 2021). Abduction is often used in exploratory research or in cases where existing theories do not fully explain a particular phenomenon to develop new theories or hypotheses. This process may involve identifying patterns, anomalies, or inconsistencies in data and to create starting points for generating possible explanations. Abduction was proposed by Charles Sanders Peirce (Osei-Bryson and Ngwenyama 2011).

Let us see an example of abduction to generate hunches in the field of psychology: A researcher might observe that a certain group of people with a particular set of characteristics tend to exhibit a specific behavior. Based on this observation, the researcher might generate a hypothesis that explains why this behavior is occurring, even though they may not have all the evidence to support it.

For instance, a psychologist who is studying anxiety may observe that people who have a history of trauma tend to exhibit higher levels of anxiety. Based on this observation, the researcher might generate a hypothesis that there is a relationship between trauma and anxiety, suggesting that anxiety serves as a coping mechanism for the trauma, even though there is insufficient evidence to support this hypothesis.

In this study, I conducted qualitative interviews with experts or elite informants in the field of intelligent process automation, including executives (head of IPA practice), program managers, process excellence consultants, programmers, and business analysts. Following are the high-level questions that were posed to the elite informants:

- ***What is your role in the IPA project?***
- ***How do you define the success of an IPA project?***
- ***What outcomes of success in IPA projects are measured?***
- ***What do you think are the predictors or critical success factors of IPA implementation?***

From stage 1, I identified the inputs that were used for the tree; in this case, I identified eleven independent variables and three dependent variables.

The second stage of the proposed methodology, as shown in Figure 3, called induction. Induction is a type of reasoning used in research that involves starting with specific observations or data and working toward a general theory or hypothesis. In other words, induction involves using observed data to develop a general understanding of a phenomenon, rather than starting with a pre-existing theory or hypothesis.

Induction is often used in exploratory research, where the goal is to generate innovative ideas or theories that can help explain the observed data or phenomenon. This process may involve identifying patterns or trends in data and using these to generate new hypotheses or theories. In this study, the data of 176 intelligent process automation projects with eleven independent variables and four dependent variables were induced into decision trees. Decision tree induction is a machine learning technique that involves constructing a decision tree from a set of training data. A decision tree is a graphical representation of a series of decisions and their possible outcomes, like a flowchart.

In decision tree induction, training data are used to build a decision tree that can then be used to make predictions or decisions based on new data. The decision tree is constructed by recursively partitioning the data based on the values of different features or variables and using these partitions to create decision nodes in the tree.

At each decision node, the algorithm selects a feature or variable that splits the data into two or more subsets based on the values of that feature. This process is repeated for each subset until a stopping criterion is met, such as reaching a maximum depth or a minimum number of samples in a node.

Once the decision tree is built, it can be used to predict the outcome of new data by following the path through the tree based on the values of the features in the new data. Each path leads to a leaf node, which corresponds to a specific outcome (I considered 3 outcomes) based on the literature review and elite informant interviews.

Decision tree induction is a powerful and widely used machine learning technique, particularly in classification problems where the goal is to predict the class of a given sample. It is relatively easy to interpret and visualize the decision tree, which can clarify the relationships between different features and the predicted outcomes.

Let us see an example of decision tree induction in the field of medicine: A researcher might be interested in developing a decision tree to predict the risk of heart disease based on various risk factors such as age, blood pressure, cholesterol levels, and family history of heart disease.

However, decision trees can also be prone to overfitting and can be sensitive to the specific ordering of features used in tree construction; to avoid this, I used tree pruning techniques to find the best representative tree with higher prediction accuracy. In stage 2, I identified the best representative trees for each of the three dependent variables (FTE reduction process efficiency and accuracy) and the impacting independent variables.

The third stage of the proposed research methodology, as shown in Figure 3, is again abduction; however, in this stage, I specifically studied the sensemaking process (Osei-Bryson and Ngwenyama 2011).

In this study, sensemaking refers to the process of interpreting and making sense of complex or ambiguous data to generate insights and understanding. This process in this study involves analyzing qualitative data such as the decision trees derived in stage 2 comprising elite informant interviews, prior literature review, and expert opinions and is based on study of 176 real-time intelligent process automation projects. This allows narrowing down the data to identify 12 independent variables or predictors used to induce the decision trees and explain the hidden phenomenon impacting the 3 dependent variables or outcomes.

Sensemaking can be particularly useful in research that involves complex or ambiguous phenomena, where existing theories or explanations may not fully capture the nuances of the data. It can also be useful in research that involves multiple perspectives or viewpoints, where the sensemaking process can help identify commonalities or differences between different participants or groups.

Let us see an example of sensemaking in the field of social sciences: A researcher might be interested in understanding the factors that contribute to job satisfaction among employees in a particular industry. The researcher could collect data through surveys or interviews with employees and use sensemaking techniques to analyze the data.

In stage 3, I will be able to identify context-specific rules and propositions using the derived results of decision trees.

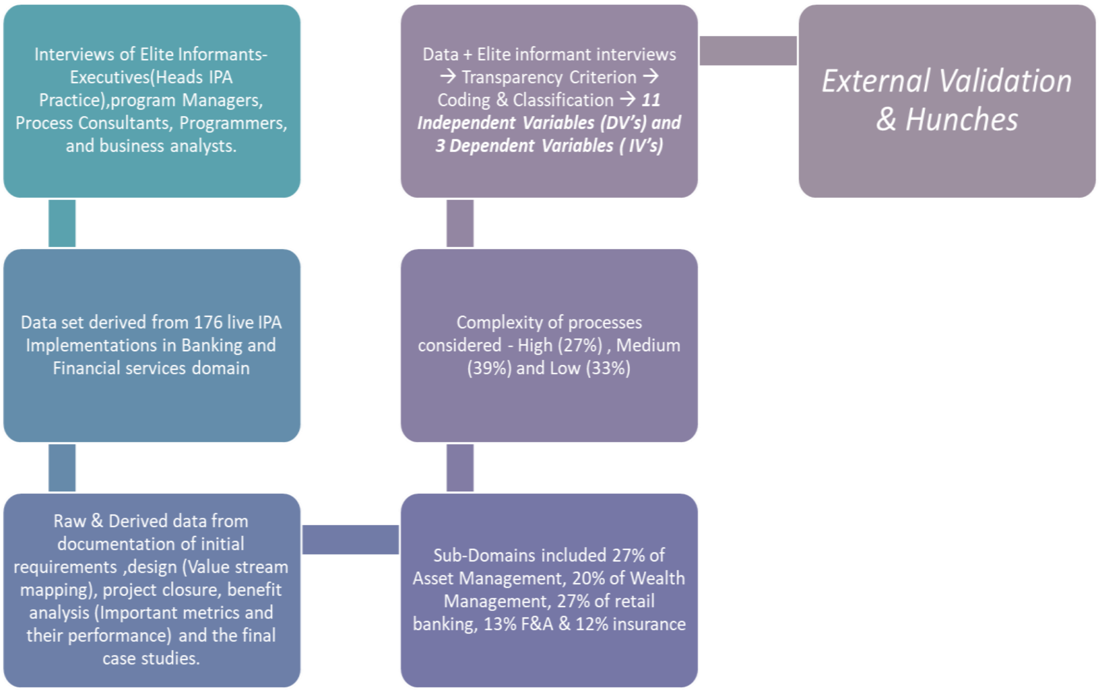

Figure 4: Data, Coding, and Classification

Figure 4 describes the process where 176 intelligent process automation implementations in the banking and financial services sector were studied. The raw data were derived from the initial requirements (including process steps and rules of automation), design (value stream mapping), project closure (mapping of initial requirements to the outcomes of automation implementation), benefit analysis (important metrics and their performance), and final case studies.

4 ABDUCTION OF PREDICTORS OF IPA SUCCESS

Through abduction, multiple theories were introduced to unfold different predictors at each level. These levels and predictors are derived from elite informant interviews, which comprised IPA practitioners, technology leaders, program managers, and business owners. Then, these predictors were categorized into four distinct levels, which have an impact on intelligent process automation success based on the data obtained from the 176 IPA implementations and elite informant interviews.

4.1 Governance-Level Predictors

Intelligent process automation (IPA) is a powerful tool for streamlining business processes and increasing efficiency. However, the implementation of IPA must be governed by appropriate guidelines (Hofmann et al., 2020, Kedziora and Penttinen, 2021) to ensure that it is used effectively. From the data and as explained by (Kedziora and Penttinen, 2021), there are different governance decisions that need to be considered before organizations embark on an IPA journey. Three predictors considered under governance levels are automation approach, automation execution, and build vs buy decision.

4.1.1 Automation Approach

When implementing intelligent process automation, one of the most important predictors under the governance level is automation Approach. In the context of this study, the automation approach is defined as the process of identifying automation opportunities and prioritize and develop automation and implementation strategies, which is an important aspect of IPA governance. The automation approach can be classified into two ways—top-down and bottom-up.

Top-down intelligent process automation (IPA) is an approach of automating business processes starting from the highest level of the process and then breaking it down into smaller tasks that can be automated using software robots or bots (Cooper et al., 2019, Kedziora et al., 2021, Naqvi and Munoz, 2020). Some of the earlier studies have discussed the importance of the top-down approach.

Bottom-up intelligent process automation (IPA) is an approach that involves starting from the task level and gradually building up to automate the entire process.

Bottom-up IPA can be a more gradual approach to automation as it allows organizations to start with small tasks and gradually build up to automate entire processes. This approach can also be useful when automating processes that are complex or involve many manual tasks (Viehhauser, 2020, Syed et al., 2020).

4.1.2 Automation Execution

In most organizations, the execution of IPA is carried out by a team that includes a combination of or individual business users (*"citizens"*) and technical experts. The specific roles and responsibilities within this team may vary depending on the organization and the nature of the IPA initiative. Automation execution is an important predictor of IPA success and is part of the governance level as it must be decided before starting the IPA journey. ***In general, the execution of the process can be triggered in two ways: citizen automation and technology-driven automation.***

Citizen intelligent process automation (IPA) is a type of IPA that is designed to be used by nontechnical business users, or "citizens," rather than dedicated IT or development teams.

Citizen IPA tools are designed to be easy to use and require little to no coding or programming knowledge (Kotsuka et al., 2019, Gorwa and Guilbeault, 2020).

With citizen IPA, business users can automate routine tasks and processes on their own, without relying on IT or development teams. This can help reduce the burden on these teams and increase productivity within the organization.

Citizen IPA has become increasingly popular in recent years as more organizations look to democratize automation and empower business users to take control of their own processes. However, it is important to note that citizen IPA tools are not a replacement for IT or development teams. Rather, they should be used in conjunction with these teams to ensure that the automation is scalable, secure, and aligned with the organization's overall technology strategy.

Technology-driven intelligent process automation (IPA) is typically referred to as "enterprise IPA" or "centralized IPA." This is a type of intelligent process automation (IPA) that is implemented and managed by dedicated IT or development teams within an organization.

Enterprise IPA is typically designed to automate more complex and mission-critical business processes that require a higher level of customization, security, and scalability. Unlike citizen IPA, which is often used to automate simple, routine tasks, enterprise IPA is used to automate more sophisticated and multi-step processes.

Overall, enterprise IPA and citizen IPA represent two different approaches to automation that are both valuable. While citizen IPA is designed to empower nontechnical users to automate routine tasks and processes, enterprise IPA is designed to tackle more complex and critical business processes that require a higher level of technical expertise and customization.

4.1.3 Build vs Buy Decision

When it comes to implementing intelligent process automation (IPA), organizations should decide whether to build their own IPA solution or purchase an off-the-shelf product from a vendor. This is known as the "build vs buy" decision.

The decision of whether to build or buy IPA depends on a variety of factors including the organization's needs, available resources, and the level of technical expertise within the organization. Some of the key considerations when making the build vs buy decision are cost, speed to market, return on investment, control on business processes, etc. (Viale and Zouari, 2020).

4.2 Process-Level Predictors

It was observed from the data that there are several predictors that are part of the business process level. These predictors have a significant impact on IPA success depending on domain (in this case, banking and financial services) categories, business processes within the domain category, and their complexity.

The three predictors under the process level are the business process domain, key process, and their complexity.

4.2.1 Domain Category

Since the domain of study is restricted to banking and financial services, various subdomains were considered as predictors of intelligent process automation (IPA) success (Oshri and Plugge 2022).

4.2.2 Key Processes

Key processes are the specific processes that are considered under the domain category to understand the insights based on the outcomes of IPA success (Thekkethil et al., 2021, Kajrolkar et al., 2021).

4.2.3 Complexity

The complexity of a business process can vary depending on the nature and scope of the process being defined. Some factors that can contribute to the complexity of business process include process scope, process interdependencies, and data complexity (Jovanović et al., 2018).

4.3 Technology-Level Predictors

When implementing IPA, it is important to consider various technology-related factors to ensure the success of the project. Some technological considerations while implementing IPA are technology architecture (scalability), interoperability (integration with other systems), technology compatibility, security, performance, and monitoring. In this study, three important technology-level predictors were considered: technology architecture, artificial intelligence, and interoperability (Auth et al., 2019, Penttinen et al., 2018, Tilson et al., 2010, Benbya et al., 2020, Torkhani et al., 2019).

4.3.1 Technology Architecture

The intelligent process automation (IPA) architecture refers to the design and structure of software that enables automation of business processes. It can vary depending on the IPA tool being used and the needs of the organization. It was observed from the study data that the technology architecture can be divided into two types: stand-alone and distributed.

Stand-alone intelligent process automation (IPA) refers to an IPA implementation that is independent of other automation or IT systems within an organization. In this type of implementation, IPA software operates in isolation, without being integrated with other enterprise applications or systems (Taulli, 2020).

Distributed intelligent process automation (dIPA) refers to an IPA architecture that allows software robots to be deployed across multiple locations or machines in a network, enabling the automation of business processes across distributed environments (Mendling et al., 2018) (Osmundsen et al., 2019, Seilonen et al., 2003, Mohanty and Vyas, 2018).

The architecture provides scalability, reliability, and improved performance, making it an ideal solution for organizations looking to automate complex business processes across their distributed networks (Miers et al., 2019).

4.3.2 Artificial Intelligence (AI)

Artificial intelligence (AI) is increasingly being used in intelligent process automation (IPA) to enhance the automation of business processes. AI can enable IPA bots to analyze data, make decisions, and learn from data, leading to more intelligent automation of complex business processes. Ways that AI is being used in IPA include document processing using OCR and image processing (Chung and Lee, 2018, LASSO-RODRIGUEZ and WINKLER, 2020).

4.3.3 Interoperability

Interoperability in intelligent process automation (IPA) refers to the ability of different IPA systems to work together seamlessly, allowing them to share data and processes across different platforms. Interoperability is becoming increasingly important as IPA systems become more widespread, and organizations seek to automate more complex business processes (Oshri and Plugge, 2022).

4.4 Complexity-Level Predictors

While intelligent process automation (IPA) is designed to automate routine, repetitive tasks, the implementation and maintenance of IPA systems can still involve complexity. Two predictors identified from studied data are coding feature and automation type (Agostinelli et al., 2019, Axmann and Harmoko, 2020) at the complexity level, which will have a significant impact on the outcome of IPA implementation.

4.4.1 Coding Feature

The coding feature in IPA refers to visual workflows that allow users to drag and drop pre-built components to build automation processes. These visual workflows provide an intuitive interface that allows users to automate processes without writing any code.

Some IPA platforms offer coding capabilities for more advanced users or for specific use cases. For example, some platforms allow developers to write custom code. This may be necessary for automating complex processes that cannot be easily accomplished with visual workflow tools (Luo et al., 2021, Agostinelli et al., 2020).

In addition, some IPA platforms allow users to integrate with APIs or web services using code. This can enable users to connect with other software systems or automate web-based processes that require more advanced scripting.

4.4.2 Automation Type

There are three main types of intelligent process automation (IPA), which are differentiated by the degree of human involvement in the automation process:

1. Attended automation: Attended automation involves the use of software robots that collaborate with human employees to automate specific tasks. These robots are typically deployed on the user's computer and are triggered by user actions, such as clicking on a button or completing a form. The robot can then take over specific tasks within the workflow, such as data entry, validation, or processing, and can be programmed to provide guidance and support to the user as needed.
2. Unattended automation: Unattended automation involves the use of software robots that work independently of human employees to automate entire business processes. These robots are typically deployed on a server or virtual machine and are programmed to run at specific times or in response to specific events. Unattended automation is useful for automating routine, repetitive tasks that do not require human intervention.
3. Hybrid automation: Hybrid automation is a combination of attended and unattended automation and is typically used for more complex business processes. Hybrid automation involves the use of software robots that can work both with and without human involvement and can switch between attended and unattended modes as needed. This allows organizations to automate more complex workflows that involve both routine, repetitive tasks and more complex decision-making processes.

Overall, the choice of the automation type will depend on the specific use case and the degree of human involvement required in the automation process (Hofmann et al., 2020, Choi et al., 2021, Berente et al., 2021a).

Table 3 summarizes the predictors of IPA success from the process of abduction as described in stage 1 of the research methodology.

Table 3: Theoretical Level and Predictors of IPA Success

Theoretical level/predictor	Definition
Governance level	
Automation approach	Process of identifying and defining automation opportunities and prioritizing and developing automation and implementation strategies for IPA. This indicates who is responsible for the automation approach
Automation execution	This indicates who executes/triggers IPA or who triggers automation
Build vs buy	This indicates whether to build IPA or to buy an off-the-shelf product
Business process level	
Domain category	Domain category of the business processes (i.e., retail, capital markets, and cards)
Key processes	Processes under the domain (e.g., KYC, onboarding, and cash management)
Intricacy	Number of steps involved in automation and systems in integration
Technology Level	
Technology architecture	The architecture chosen to implement IPA
Artificial intelligence	The level of AI required driving automation of processes
Interoperability	Ability of different IPA systems to work together seamlessly
Complexity	
Coding feature	Extent of coding required to automate a process through IPA
Automation type	The process type selected and designed for IPA implementation

4.5 Conclusion

This chapter defined how the studied data and elite informant interviews were used to arrive at the theoretical levels and predictors or critical success factors of IPA. Each of the predictor along with the prior references from literature reviews was defined. The next chapter will describe the research design, by providing the context and explaining about the data sources and three-stage research methodology.

5 DATA AND MEASURES

In this study, the research context was a Fortune 200 organization, from where the data set was derived. This chapter describes why banking and financial services was chosen as the context of the study and defines the four measures of outcome and the rationale behind coding of success for these measures, followed by measures of eleven predictors and the rationale for coding these predictors.

5.1 Research Context

The data were collected from a Fortune 200 US multi-national information technology services and consulting company. The company has an intelligent process automation practice worth over $200 million, which has been ranked a leader in the 2022 Everest Group Peak Matrix for intelligent automation providers. The IPA group within the company engaged in delivering more than 2,000 projects between 2019 and 2021 across various domains such as banking and financial services, media and communications, retail, healthcare, and life sciences, where intelligent process automation was implemented.

5.2 Data

In total, 700 projects were considered for the study, but due to numerous subdomains, complexity of processes, and commonality of processes across geographic locations, the analysis was limited to 176 IPA implementation projects in the ***banking and financial services*** domain to identify the success factors of intelligent process automation. These elite informant interviews were restricted to the intelligent process implementations from the banking and financial services domain to prevent interdomain heterogeneity and to control the interdomain variance.

Various business processes can be automated using intelligent process automation across subdomains such as "retail banking," "wealth management," "asset management," "finance and accounts," and "insurance." The business processes that were considered under banking and financial services include 27% of asset management, 20% of wealth management, 27% of retail banking, 13% F&A, and 12% of insurance, where intelligent process automation has been implemented. These statistics were in accordance with the number of overall processes across subdomains.

Of the total number of projects studied, 27% of the projects considered were of high complexity, 39% of were of medium complex, and 33% were of low complex. These make the data homogeneous and avoid selection bias. Thus, focusing on banking and financial services serves as a key sampling criterion in this study.

In total, more than 100 documents were studied, and five elite informant interviews were conducted for this study.

The combination of elite informant interviews and the data derived from the 176 intelligent process automation interviews were further classified and coded to arrive at eleven independent variables and three dependent variables.

5.3 Measure of Outcome of Interest: IPA Success

The purpose of this study was to inductively build a multi-level theory for identifying the factors impacting the success of intelligent process automation (IPA) by rigorously analyzing the implemented IPA projects. Several previous studies have focused on evaluating why some programs have achieved significant value, while others have fallen below expectations. However, the literature (Lacity et al., 2021) provides action principles to guide leaders through their intelligent automation journey. The following three outcomes have been found to be critical for the successful implementation of intelligent process automation.

Considering prior research, expert opinions, elite informant interviews, and ROI analysis of the 176 IPA implementations in stage 1, as shown in Figure 3, three outcome variables, namely, *"full-time equivalent (FTE) reduction, "process efficiency," and "accuracy"* (Asatiani and Penttinen, 2016, Benbya et al., 2021), were found to be the most significant to measure the success factors of intelligent process automation (IPA).

Based on the expert opinions and elite informant interviews of several technologists who specializes in intelligent process automation (IPA), literature review, and the analysis of the 176 IPA projects, the following success predictors for intelligent process automation were identified, and this is the result of stage 1 of the research process, as depicted in Figure 3.

5.3.1 Full-Time Equivalent (FTE) Reduction

In the context of intelligent process automation (IPA), FTE reduction refers to the number of full-time employees that can be replaced by software robots that automate repetitive and rules-based tasks (Asatiani and Penttinen, 2016, Lacity and Willcocks, 2021, Willcocks et al., 2017, Wewerka and Reichert, 2020).

IPA is designed to streamline business processes and reduce the time and effort required to complete tasks. By automating tasks that were previously performed by human workers, IPA can significantly reduce the number of FTEs required to complete a task or process. This is because IPA can work continuously, without breaks or downtime, and can perform tasks at a much faster rate than humans.

FTE reduction value when automating a process through IPA is assigned as "high" when FTE reduction is more than 70%, "medium" when between 50% and 70%, and "low" when lower than 50%. The *higher the FTE reduction, the higher the success probability of intelligent process automation.*

5.3.2 Average Handling Time (AHT)

Average handling time (AHT) is a metric used to measure the time taken to complete a particular task or process. In the context of intelligent process automation, AHT refers to the time taken for a software robot to complete a task or process.

AHT is an important metric for organizations that are using IPA to automate their processes because it helps measure the efficiency and effectiveness of their automation efforts. By reducing AHT, organizations can improve their operational efficiency, reduce costs, and increase customer satisfaction.

AHT is highly co-related with FTE reduction and is categorized as "high," "medium," and "low." *The lower the AHT, the higher the success of intelligent process automation.*

5.3.3 Process Efficiency

Process efficiency is critical for intelligent process automation as it enables organizations to reduce costs, improve productivity, and increase accuracy. In the context of IPA, process efficiency refers to the ability of software robots to perform tasks and processes quickly, accurately, and without errors. Process efficiency signifies an optimal (in most of the cases, the fastest or the cheapest) way of operating processes (Riemer and Peter, 2020, Carden et al., 2019, Berente et al., 2021b, Asatiani et al., 2020) and can be measured by the amount of effort required to achieve a business outcome. Process efficiency is considered to be "high" when the automation is more than

90%, "medium" when between 70% and 90%, and "low" when less than 70%. The *higher the process efficiency, the higher the success of intelligent process automation.*

5.3.4 Accuracy

Accuracy is a critical factor for successful robotic process automation (RPA) implementation. Accuracy refers to the ability of software robots to perform tasks and processes with a high degree of precision and without errors. Inaccurate automation can lead to costly mistakes, lost time, and reduced efficiency.

Accuracy is another parameter that indicates the extent to which a process is automated and how successful it is when the process is run several times (Asatiani et al., 2020, Riemer and Peter, 2020, Benbya et al., 2020). Accuracy is considered to be "high" when it is above 95%, "medium" when between 80% and 95%, and "low" when less than 80%. The *higher the accuracy, the higher the success of intelligent process automation.*

Other automation success variables that were studied were *usability* categorized as "high," "medium," and "low"; *payback time* categorized as "fast," "medium," and "slow"; and *repeatability* of the process categorized as "yes" and "no."

Table 4 depicts the measures of outcome of interest for IPA success.

Table 4: Measures for Outcome of Interest

Success predictors		
Predictor	Category	Rationale
Full-time equivalent (FTE) reduction	High	>70%
	Medium	50% to 70%
	Low	<50%
Process efficiency	High	>90%
	Medium	70% to 90%
	Low	<70%
Accuracy	High	>95%

	Medium	80% to 95%
	Low	<80%

5.4 Measures of Governance-Level Predictors

Automation of any process depends on how it is governed; in other words, who are the key decision-makers and stakeholders who drive the intelligent automation initiatives. Automation of a business process needs a fair bit of governance to ensure the success even before organizations embark on the automation journey.

5.4.1 Automation Approach

Automation approach is defined in two ways: top-down and bottom-up. Top-down means that the automation mandate is part of the organization mandate of digitizing the enterprise and its business process (Cooper et al., 2019, Kedziora et al., 2021, Naqvi and Munoz, 2020), whereas bottom-up means low-level process automation implemented by business users implementing the tasks.

5.4.2 Automation Execution

Who executes the automation forms the part of automation execution. In general, the execution of a process can be triggered in two ways: citizen automation and technology-driven automation.

When the business user does the execution, it is called "citizen automation" (Kotsuka et al., 2019, Gorwa and Guilbeault, 2020), and when the execution is done centrally through a product or technology group, it is known as "technology-driven automation."

5.4.3 Build vs Buy

A decision should be made on build vs buy, that is, whether to execute the automation through an internally developed tool or through a readily available product tool, or sometimes, a combination of both is used by customizing the product; hence, it categorized as "buy," "build," or "both."

Table 5: Measures of Governance-Level Predictors

Governance-level predictors		
Predictor	Category	Rationale
Automation approach	Top-down	Executive-driven
	Bottom-up	Process owner-driven
Automation execution	Citizen bot	Business-controlled
	Technology-driven	Technology-controlled
Build vs buy	Buy	Off-the-shelf product
	Build	Bespoke development
	Both	Combination

5.5 Measures of Process-Level Predictors

5.5.1 Domain Category

The study of interest here is banking and financial services, and the aim of this study was to understand the IPA success for the processes under the BFS *domain category* such as "asset management," "wealth management," "retail banking," "finance and accounts," and "insurance" (Oshri and Plugge, 2022).

5.5.2 Key Processes

The predictor *key processes* lists down an IPA implementation at a business process level under each domain category (Thekkethil et al., 2021, Kajrolkar et al., 2021).

5.5.3 Complexity

Each of the 176 key business processes analyzed were associated with varying degrees of *complexity* (Timbadia et al., 2020, Berente et al., 2021a, Asatiani and Penttinen, 2016, LASSO-RODRIGUEZ and WINKLER, 2020) and categorized as high, medium, or low. This categorization was based on discussion with several industry experts and elite informant interviews from stage 1 of the research methodology, as depicted in Figure 3. A business process with more than one hundred steps/rules to be automated that is highly distributed/interoperable is assigned a high value, a business process with 50-100 steps/rules with a distributed/interoperable flow is assigned a medium value, and a business process with less than 50 steps with no interoperability is assigned a low value.

Table 6: Measures of Process-Level Predictors

Process-level predictors		
Predictor	Category	Rationale
Domain category	Right domain	Retail, asset, etc.
Key processes	Right process	Onboarding and accounting, etc.
Intricacy	High	>100 steps
	Medium	50-100 steps
	Low	<50 Steps

5.6 Measures of Technology-Level Predictors

Implementation of IPA can be achieved using an off-the-shelf product, or it can be custom-built through bespoke development, or it can be a combination. *Technology selection* is an important criterion when analyzing the success of IPA.

5.6.1 Architecture

The technology architecture plays a key role based on the process to be automate and has an impact on the overall success. The architecture is categorized as "standalone" or "distributed." Standalone automation primarily automates mundane and repetitive tasks done by humans but cannot automate end-to-end processes with assistance. Distributed automation is more holistic involving AI/ML and strong coding-based architectures to enable full automation across multiple business lines within an enterprise. Recent progress in artificial intelligence, machine learning, cryptography, and cloud-based distributed systems have provided new technologies for distributed intelligent process automation integrating several internal and external systems (Mendling et al., 2018), thereby providing an control on end-to-end process view and automation (Osmundsen et al., 2019, Seilonen et al., 2003, Mohanty and Vyas, 2018).

5.6.2 Artificial Intelligence (AI)

Automation of a process can be done using readily available product RPA tools or in combination with artificial intelligence (Chung and Lee, 2018, LASSO-RODRIGUEZ and WINKLER, 2020) to improve the automation success, so *AI* plays a significant role in automation. In my data analysis it is categorized as high or low

5.6.3 Interoperability.

Interoperability defines if process is cut across multiple systems to execute the activities/tasks while integrating with other systems (Oshri and Plugge, 2022). It is categorized as "yes" if there is an interaction and "no" if there is no interaction.

Table 7: Measures of Technology-Level Predictors

Technology-Level Predictors		
Predictor	Category	Rationale
Architecture	Standalone	Task-level automation
	Distributed	Process-level automation
AI	Yes	Uses AI/ML
	No	Does not use AI
Interoperability	Yes	Cuts across systems
	No	Single system

5.7 Measures of Complexity-Level Predictors

While intelligent process automation (IPA) is designed to automate routine, repetitive tasks, the implementation and maintenance of IPA systems can still involve complexity and can be explained in terms of coding feature that is involved in automation and the type of automation.

5.7.1 Coding Feature

Automation complexity has several aspects, and the most important aspect is *coding* (Luo et al., 2021, Agostinelli et al., 2020). It is categorized into "strong," "average," and "low" coding;

for example, when there is low of bespoke code written to automate the process, it is categorized as strong coding.

5.7.2 Automation Type

The *automation type* depends on the extent of the manual intervention required when a process is automated. It is categorized as "unattended" when there is no manual intervention, "attended" when there are more steps of manual intervention than what robots accomplish, and "hybrid" when majority of automation is accomplished by robots and some critical tasks need human intervention to make the process successful (Hofmann et al., 2020, Choi et al., 2021, Berente et al., 2021a)

Table 8: Measures of Complexity-Level Predictors

Complexity-Level Predictors		
Predictor	Category	Rationale
Coding	Strong	Product +custom Code
	Average	Product +configuration
	Low	Low code
Automation type	Unattended	Completely automated
	Hybrid	Humans + bots
	Attended	Humans

:

5.8 Descriptive Statistics

The table below shows the frequency table for all the independent and dependent variables studied in this thesis.

Table 9: Frequency table

Variable	Categories	Frequency	Percentage
Complexity	High	48	27.27%
	Low	57	32.39%
	Medium	71	40.34%
Architecture	Distributed	49	27.84%
	Stand Alone	127	72.16%
AI/ML	No	43	86.36%
	Yes	133	13.64%
Interoperability	No	43	24.43%
	Yes	133	75.57%
Coding Feature	Average Coding	55	31.25%
	Low Coding	91	51.70%
	Strong Coding	30	17.05%
Automation Type	Attended	29	16.48%
	Hybrid	77	43.75%
	Unattended	70	39.77%
Automation Approach	Bottom Up	107	60.80%
	Top Down	69	39.20%
Automation Execution	Citizen Bots	151	85.80%
	Tech Driven Bots	25	14.20%
Build Vs Buy	Build	16	9.09 %
	Buy	156	88.64 %
	Both	4	2.27 %
Accuracy	High	104	59.09%
	Low	67	38.07%
	Medium	5	2.84%
Process Efficiency	High	51	28.98%
	Low	63	35.80%
	Medium	62	35.23%
FTE Reduction	High	85	48.30%
	Low	50	28.41%
	Medium	41	23.30%

Below is the correlation table for all the independent and dependent variables studied in this thesis.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) FTE Reduction	1.000											
(2) Process Efficiency	0.659*	1.000										
(3) Accuracy	0.447*	0.305*	1.000									
(4) Automation Approach	0.337*	0.193*	0.417*	1.000								
(5) Automation Execution	0.019	0.067	0.038	-.293*	1.000							
(6) Automation Type	0.550*	0.574*	0.434*	0.099	0.384*	1.000						
(7) Complexity	-.157*	-.199*	-0.116	0.052	-.280*	-.309*	1.000					
(8) Architecture	0.175*	0.295*	0.096	0.176*	0.256*	0.327*	-.337*	1.000				
(9) AI/ML	0.063	-0.028	-0.070	0.251*	-.645*	-.362*	0.392*	-.418*	1.000			
(10) Interoperability	-0.146	-.197*	0.014	-.213*	0.034	-.167*	0.220*	-.235*	0.072	1.000		
(11) Coding Feature	-0.123	-.152*	-0.142	-0.107	-.425*	-.283*	0.527*	-.437*	0.578*	0.142	1.000	
(12) Build Vs Buy	0.002	0.067	0.016	-.260*	0.801*	0.426*	-.375*	0.345*	-.821*	-.018	-533*	1.000

**** p<0.01, ** p<0.05, * p<0.1*

5.9 Conclusion

This chapter has provided the research context on the source of data that are derived from live IPA implementations in the banking and financial services domain and elite informant interviews and described the data collection process to form the external validation or hunches to arrive at eleven independent variables and four dependent variables or outcomes of success for IPA implementation. Finally, the three-stage research methodology of abduction, induction, and abduction has been explained. The next chapter will define the measures encompassing the four outcome variables and eleven independent variables and assign or code the values for the data analysis.

The chapter has discussed about the measures for outcome of interest, theoretical levels, and the eleven predictors under these levels. It also covered the mechanism of how they were assigned values and the rationale for classification of values used for the data analysis to identify the critical success factors of intelligent process automation (IPA). The next chapter will describe the decision tree induction process from the data abducted from the 176 live IPA implementations and the elite informant interviews.

6 INDUCTION OF DECISION TREES

Insights into and patterns in data are derived using a data-first methodology of tree induction (Quinlan, 1986b, Quinlan, 1990). Decision tree induction is used to construct rules that demystify the information from a data set, the interpretation of which helps managers take informed and data-driven decisions (Boonstra, 2003) and helps construct the best representative structure to solve complex scenarios (Alter, 1978). This methodology helps in informed decision-making, which is critical to leaders and executives when embarking on transformational initiatives (Boonstra, 2003, Karhade et al., 2015)(Langley et al., 1995, Counihan et al., 2002, Alter, 1978)(Lin et al., 2017).

This analytical approach involves a series of if-then statements derived from the tree structure, making it easy for stakeholders to understand factor-related outcomes. Decision trees clearly show the relationship between important predictors and outcomes based on the actual results, rather than an empirical forecast (Langley et al., 1995, Markus et al., 2002). Decision tree induction helps researchers derive a considerable number of hypotheses rapidly and generate meaningful insights (Osei-Bryson and Ngwenyama, 2011). A few broad assumptions about the data and their distribution are applied to the decision tree algorithm (I discuss C4.5) to increase its applicability and generalization (Quinlan, 1986b)(Quinlan, 1990)).

Organizational environments and decisions are supported by using the decision tree induction methodology based on what should happen, and not how they are forecasted to happen (Markus et al., 2002, March, 1994). The decision tree framework has applications (Drazin and Van de Ven, 1985, Pomerol et al., 2002)in numerous fields, particularly Fintech (Lagna and Ravishankar, 2022), IT portfolio management (Counihan et al., 2002, Otim and Grover, 2012)], persona-based human development decision-making (Bailey and Ngwenyama, 2014), financial decision-making (Tessmer et al., 1993), healthcare decision-making (Lin et al., 2017), and, more recently, human-artificial intelligence (AI) augmentation due to extensive automation (Sturm et al., 2021) (Wang et al., 2021).

6.1 Data Partitioning in Decision Tree Induction

Data partitioning is the key to generating theory through tree induction, especially when there is a large amount of observational data, to arrive at decisions with propensity, which provides finer insights into and easy understandability and use of data (Yahav et al., 2016). In data partitioning, the sample training data set is divided into smaller subsets with the growth of the tree, so that the most relevant attribute is identified efficiently, establishing generality and enabling accurate prediction of unseen data to generate theory (Tessmer et al., 1993).

Decision tree partition is examined through n-fold validation, where data sets are divided into n partitions, n-1 partitions are used as the training subsample, and one partition (or fold) is used for validation. In this analysis, 10-fold validation was used and is one of the most accurate and popular testing modes for building theory using decision tree induction (Hibbeln et al., 2017) (Karhade et al., 2015). WEKA data mining software was used, and measures were taken to avoid overfitting of data and to achieve higher prediction accuracy (Hall et al., 2009).

There are two steps in tree induction following data partitioning: first, the C4.5 induction algorithm is applied on training data to build a decision tree (Quinlan, 1986b) (Quinlan, 1990)], and second, the constructed tree is pruned by performing various computational experiments to identify the tacit structure of data and signifies the robustness of knowledge discovery. WEKA was used for data partitioning, growing, and pruning trees (Hall et al., 2009).

The C4.5 algorithm evaluates the goodness of fit of the data for generating maximum information from the data sets and satisfactorily manages common issues that arise in decision tree construction (Quinlan, 2014). Tree induction iteratively groups observations (i.e., intelligent process automation implementations in banking and financial services, in this case) that have similar attributes and similar outcomes that predict the success of intelligent process automation (i.e., FTE reduction, process efficiency, and accuracy). Broadly, there are two key inputs for decision tree induction: 1) 176 intelligent process automation (IPA) implementation attributes or variables, and 2) success factors affecting IPA implementation.

Tree induction was carried out to identify information attributes with similar outcomes (in this case, FTE reduction, process efficiency, and accuracy of IPA implementation) (Quinlan, 1986b).

Using prediction accuracy alone as a criterion for choosing the best representative tree can be an overfitting trap, and to avoid the overdependency on prediction accuracy, two additional heuristics, namely, communicability (parsimoniousness) and structural consistency (stability) of the discovered knowledge, are introduced (Boonstra, 2003). In summary, the choice of the best representative tree is a combination of three heuristics: 1) prediction accuracy, 2) parsimoniousness, and 3) consistency of the tree structure, i.e., overall stability of the discovered knowledge.

It needs to be understood that the implied basic decision rationale uncovered based on the tree induction does not reflect the exact rules of decision-makers (Boonstra, 2003). Decision tree induction avoids the correlated predictors and only conveys the most informative knowledge, which has a strong impact on final decision outcomes. In this study, correlations between predictors were not reported. In the following section, data analysis is described, which includes computational experiments that enable selecting the best representative tree for implementing intelligent process automation (IPA) based on three outcome predictors 1) full-time equivalent (FTE) reduction, 2) process efficiency, and 3) accuracy.

6.2 Need for Computational Experiments

Tree pruning helps reduce the size of the tree by removing a part of the tree that has little power to classify the instances and helps provide better accuracy by reducing overfitting and noise or erroneous data (Hssina et al., 2014), thereby ensuring that rational decisions are comprehensively discovered and the training and subsamples are repeated many times. In each run, two mutually exclusive subsamples of were drawn from the studied 176 intelligent process automation implementations. The first set of subsamples is known as the training set, and the second set is known as the testing set. The training set was used to find the tacit decision rationale by using the C4.5 induction algorithm (Quinlan, 1986b), and the testing set was used to derive the predictive accuracy of the discovered decision (Boonstra, 2003).

6.3 Selecting the Best Representative Tree

In this analysis, tree selection is based on 10-fold validation as the empirical evidence (Weiss and Indurkhya, 1994) suggests that 10-fold validation is unbiased, consistent with optimal tree selection, and accurate irrespective of population distribution as it is dependent on the sample size. In 10-fold validation, the data set is divided into ten subsets and repeated ten times. In every run, one of the ten subsets is used as a test set, and the remaining nine subsets are used as the training set used for building a tree (Quinlan, 1990)1].

Decision rationale to predict the success factors of intelligent process automation from the unseen data is obtained by assessing the prediction accuracy of the tree generated from the training set.

10-fold validation is repeated at various levels of pruning, i.e., different confidence factors and communicability of the tree, by repeating computational experiments from which multiple approximations of decision rationale are derived. This is fundamental to the decision tree induction methodology so that multiple estimations for underlying decisions are available for researchers.

6.4 Three Key Heuristics

The three heuristics discussed earlier: 1) High prediction accuracy, 2) Parsimony, and 3) consistency to select the best representative decision tree, which, in turn, make the decision rationale credible.

1) High predictive accuracy: Prediction accuracy of trees induced on the training data is evaluated on a mutually disjoint validation data set. This heuristic represents a goodness-of-fit measure in terms of predicting decisions from unseen data.
2) High parsimony: The induced tree is expected to be a parsimonious approximation of the underlying decision rationale so that it can serve as an effective decision- and policy-making aid.
3) High reliability: Since the process of drawing training samples to induce trees and evaluating the predictive accuracy of induced trees on mutually disjoint validation samples is repeated several times, the robustness of the discovered knowledge was assessed. In this case, all the trees induced on the data contained the same topmost attribute, showing up reliably across these multiple iterations, thereby representing a robust approximation of the underlying decision rationale.

Thus, the trees presented in this research are credible approximations of the intelligent process automation success outcomes in terms of "FTE reduction," "process efficiency," and "accuracy".

6.5 Examining Full-Time Equivalent Reduction

FTE reduction is one of the most important automation goals and has been found to impact the success of IPA implementation (Engel et al., 2021), and the released FTEs are redeployed for the jobs that are strategic in nature (Coombs et al., 2020).

Table 2 outlines different computational experiments that were performed to identify the most suitable representative tree and facilitate the development of theories to pinpoint the critical factors impacting FTE reduction in the realm of intelligent process automation. By manipulating two primary parameters (confidence level and minimum instances at leaves) during the tree pruning process, various models were generated. In this analysis, four distinct levels (i.e., governance, process, technology, and complexity levels) were found, which encompassed a total of ten predictors to model the success factors of intelligent process automation with respect to FTE reduction.

Table 10: Computational experiments to select best representative tree for FTE reduction.

Mechanism to Detect Overfitting			**Three Key Heuristics to Choose the Best Representative Model**		
No.	Degree of pruning	Minimum instances at leaves	COMMUNICABILITY Size of tree (# of leaves)	CONSISTENCY Topmost attribute	ACCURACY Prediction Error (Validation Data)
1	High (0.25)	8	11	1.Automation type 2. Top-down/bottom-up	38%
2	Medium (0.5)	8	13	1.Automation type	39.20%

				2. Top-down/bottom-up	
3	Low (0.75)	8	13	1.Automation type 2. Top-down/bottom-up	40%
4	High (0.25)	10	11	1.Automation type 2. Top-down/bottom-up	38%
5	Medium (0.5)	10	13	1.Automation type 2. Top-down/bottom-up	38.6%
6	Low (0.75)	10	13	1.Automation type 2. Top-down/bottom-up	39.77%
7	High (0.25)	12	11	1.Automation type 2. Top-down/bottom-up	36.30%
8	**Medium (0.5)**	**12**	**11**	**1.Automation type 2. Top-down/bottom-up**	**35.70%**
9	Low (0.75)	12	11	1.Automation type 2. Top-down/bottom-up	35.70%
10	High (0.25)	14	9	1.Automation type 2. Top-down/bottom-up	36.36%
11	Medium (0.5)	14	9	1.Automation type 2. Top-down/bottom-up	34.65%
12	Low (0.75)	14	9	1.Automation type	34.65%

				2. Top-down/bottom-up	
13	High (0.25)	16	9	1.Automation type 2. Top-down/bottom-up	34.70%
14	Medium (0.5)	16	9	1.Automation type 2. Top-down/bottom-up	34.70%
15	Low (0.75)	16	9	1.Automation type 2. Top-down/bottom-up	34.70%

Computational experiments allow choosing variables affecting the success of FTE reduction implementing intelligent process automation diligently through an incremental approach for the development of theory.

6.5.1 Selecting the Best Representative Tree FTE Reduction

To make sure that decision justification is found comprehensively, a method of drawing mutually exclusive training and testing samples is repeated multiple times. In every repetition, two random mutually exclusive subsamples of intelligent process automation derived from the data set of 176 projects were drawn with FTE reduction as the outcome variable, which is one of that factors that represent the success of intelligent process automation. Of two sets, one set is used as the training set, from which implicit decision rationale is found by using the C4.5 induction algorithm (Quinlan, 1986a), and the second set is known as the testing set, which is disjoint and used to evaluate the predictive accuracy of implicit decision rationale. The decision tree induction uses 10-fold validation, last partition of which is used for validation by comparing it against the other partitions used for building the tree. Prediction accuracy of the obtained tree from the training set is assessed by applying the rationale to predict FTE reduction, which is one of the factors effecting the success of intelligent process automation from unseen data in the testing set, which is a mutually disjoint set.

Multiple calculations of inherent implicit decision rationale were performed by repeating 10-fold validations as part of computational experiments. This is part of the tree induction methodology to make sure that multiple estimates of implicit decision rationale are available to researchers.

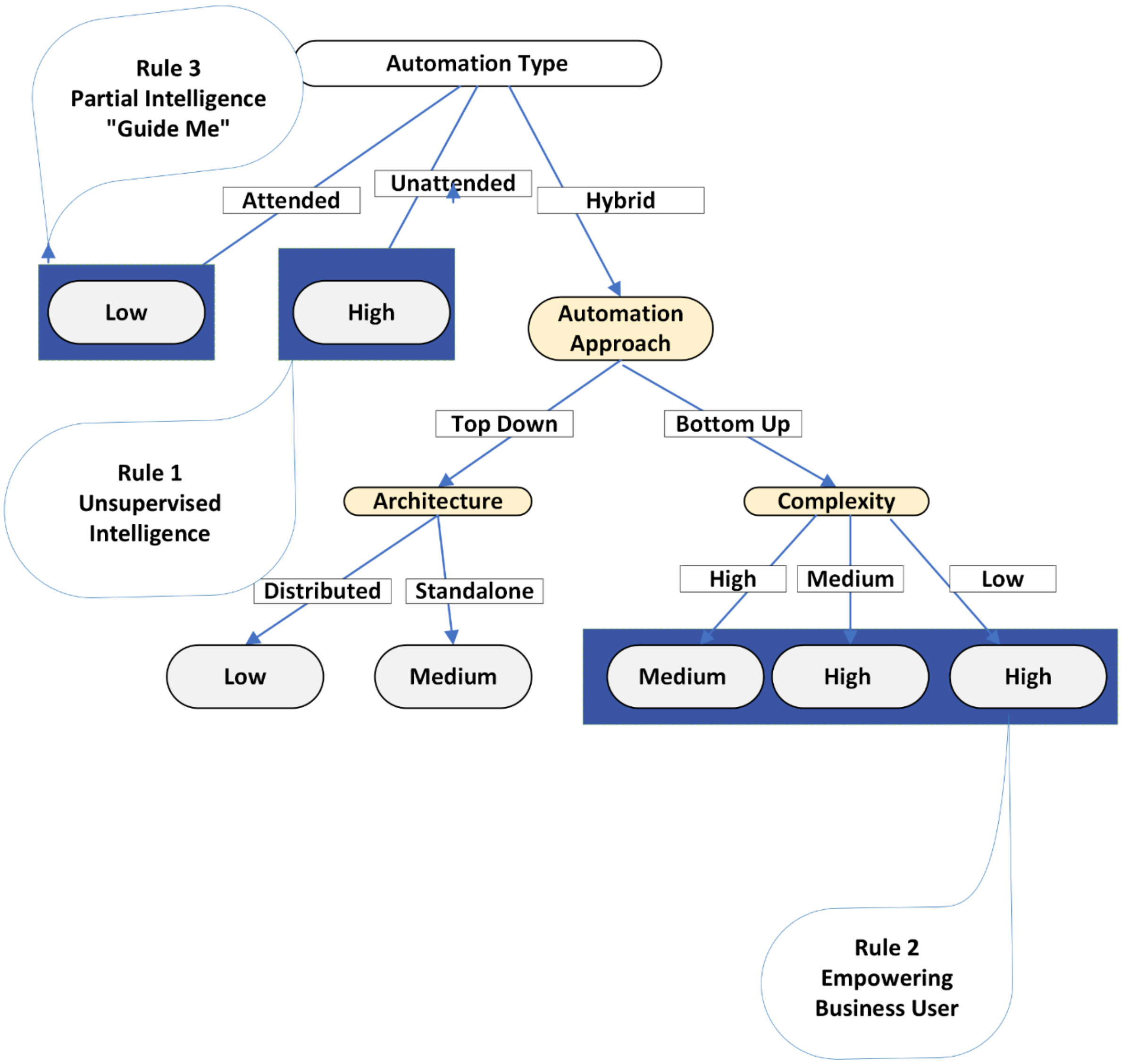

Figure 5: Tree 1: Decision Journey for FTE Reduction

Tree 8 from Table 10 depicted in Figure 5 is the best representative tree; although prediction error is marginally higher than the least prediction error, it optimally satisfies the three key heuristics of low prediction error, parsimony, and consistency of the top-level variable.

Based on the information provided, it can be inferred that the most significant factors influencing FTE reduction in the context of intelligent process automation are the "automation type" and the approach adopted for automation, i.e., "top-down/bottom-up." The success of intelligent process automation is directly proportional to the degree of FTE reduction achieved, indicating that higher FTE reduction levels correspond to greater success in this regard.

6.5.2 Robustness Check

The best representative tree, i.e., tree 8 from Table 1, was obtained through using 10-fold cross-validation. To ensure the robustness of the result, a computational experiment on the best representative tree was also performed using the percentage split test using 80%, which means that 80% of the 176 instances are trained and tested with reminder of 20% of instances. I find that both with 10-fold validation and 80% percentage split the best representative tree 8 yields same results.

6.6 Key Findings & Rules for FTE Reduction

Trees, in Figure 5, that were discovered by the tree induction C4.5 algorithm, are not the precise rules or "a written rulebook" that decision-makers implementing IPA can use. Instead, they are estimates of the inherent structure of the data. The trees yield context-specific rules that clearly articulate the emergent connections across levels to inform the IPA success factors and therefore constitute part of my multi-level theory.

All the eleven information attributes characterizing initiatives in conjunction with the final FTE reduction decision are inputs to tree induction. All information attributes discovered by tree induction to be most informative for explaining FTE reduction decisions are included in the tree as decision attributes, and the tree induction C4.5 algorithm excludes all the noninformative attributes from the tree. The most informative decision attribute is the topmost attribute in the tree. The importance of attributes decreases as we move away from the top of the tree to its leaves.

Three rules were derived in this study, as depicted in Figure 5: rules 1 and 2 are the top rules that predict the factors for high FTE reduction, resulting in successful IPA implementation, whereas rule 3 predicts factors for low FTE reduction, resulting in low IPA implementation. Rules are presented in Table 10.

Table 11: Top context-specific rules discovered from decision tree induction (FTE reduction)

No	Rule	Levels incorporated	Decision (FTE reduction)	Instances classified
1	Unsupervised intelligence *Automation type =“unattended”*	*Governance* *Process* *Technology* *Complexity*	High	80
2	Empowering business user *Automation type =“hybrid” and* *Automation approach = “bottom-up” and* *Complexity = “high” or “medium” or “low”*	*Governance* *Process* *Technology* *Complexity*	High	56
3	Partial intelligence “guide me” *Automation type = “attended”*	*Governance* *Process* *Technology* *Complexity*	Low	41

6.6.1 Rule 1: Unsupervised Intelligence

Unattended automation executes tasks of a process without any human involvement from the start to the end, where the process is mostly scheduled to start or some event triggers the process to begin. Unattended automation manages manual tasks that involves a specific pattern or a specific set of steps that are meant to be followed. From the studied 176 implementations for the outcome variable "FTE reduction," it is observed that FTE reduction is very high when the automation is unattended. This was the main classification rule extracted from the tree as it is classified as most IPA implementations (80 IPA implementations). The general form of the rule is ***when automation is unattended, FTE reduction obtained in an IPA implementation is high.***

6.6.2 Rule 2: Empowering Business User

When automation **involves some user input**, hybrid automation is recommended. With a hybrid automation model, attended automation performs the part that requires human intervention, and the rest is performed by unattended automation, and vice versa happens when unattended automation requires humans to make decisions. Bottom-up automation is **driven by business owners at the process level** who are empowered to give ideas as they have the complete knowledge of the domain and the gaps to be automated, with an all-inclusive approach that results in successful automation of a process; this kind of bottom-up approach will also make FTEs more productive. From the 176 IPA implementations, when "FTE reduction" is used as the outcome variable, when the automation type is hybrid, i.e., combination of attended and unattended automation, and when the **business process owners are involved in the decision-making of automation**, irrespective of complexity due to the inclusive approach, automation success is either high or trending toward high. This was the main classification rule extracted from the tree as it is classified most IPA implementations (56 IPA implementations). The general form of the rule is ***when automation is hybrid and driven by a bottom-up approach, FTE reduction obtained in an IPA implementation is high irrespective of process complexity.***

6.6.3 Rule 3: Partial Intelligence "Guide Me"

Attended automation involves working alongside humans and managing certain tasks within longer more complex work sequences or processes that cannot be fully automated from the start to the end. Attended automation generally occurs when there is no specific pattern identified for the business process and can only be performed by humans. From the studied 176 implementations for the outcome variable "FTE reduction," it is observed that FTE reduction is very low when automation is attended. This was the main classification rule extracted from the tree as it classified most IPA implementations (41 IPA implementations). The general form of the rule is ***when automation is attended, then the FTE reduction obtained in an IPA implementation is low.***

6.7 Examining Process Efficiency

Process efficiency is one of the key success indicators of intelligent process automation (Asatiani and Penttinen, 2016, Santos et al., 2019) and has been found to be an expected impact when considering success of intelligent process automation. Higher process efficiency leads to faster business growth post intelligent process automation.

The data set for process efficiency undergoes the same process of computational experiments, pruning through 10-fold validation with the same set of ten predictors, as discussed earlier for FTE reduction, to generate decision rationale and theory development to derive the predictors impacting the process efficiency of intelligent process automation.

Computational experiments for process efficiency are depicted in **Table 11.**

Table 12:Computational experiments to select best representative tree for process efficiency.

Mechanism to detect overfitting			Three key heuristics to choose the best representative model		
No.	Degree of pruning	Minimum instances at leaves	Communicability	Consistency	Accuracy
			Size of tree (# of leaves)	Topmost attribute	Prediction error (validation data)
1	High (0.25)	8	6	1.Automation type 2.Architecture	42.60%
2	Medium (0.5)	8	15	1.Automation type 2.Interoperability, automation execution	47%
3	Low (0.75)	8	23	1.Automation type 2.Interoperability, automation execution	46.60%
4	High (0.25)	10	6	1.Automation type 2.Architecture	46.60%
5	Medium (0.5)	10	14	1.Automation type 2.Interoperability, automation execution	47.10%
6	Low (0.75)	10	14	1.Automation type 2.Interoperability, automation execution	47.10%
7	High (0.25)	12	6	1.Automation type 2.Architecture	46.60%
8	Medium (0.5)	12	6	1.Automation type 2.Architecture	47.70%
9	Low (0.75)	12	6	1.Automation type 2.Architecture	48.30 %
10	High (0.25)	14	6	1.Automation type 2.Architecture	44.90%
11	Medium (0.5)	14	6	1.Automation type 2.Architecture	43.75
12	Low (0.75)	14	6	1.Automation type 2.Architecture	46%
13	High (0.25)	16	6	1.Automation type 2.Architecture	46%
14	Medium (0.5)	16	6	1.Automation type 2.Architecture	46%

15	Low (0.75)	16	6	1.Automation type 2.Architecture	46%

6.7.1 Selecting the Best Representative Tree "Process Efficiency"

Figure 6 depicts the best representative tree that provides a rationale indicating best predictors for success of IPA implementation in terms of process efficiency as an outcome.

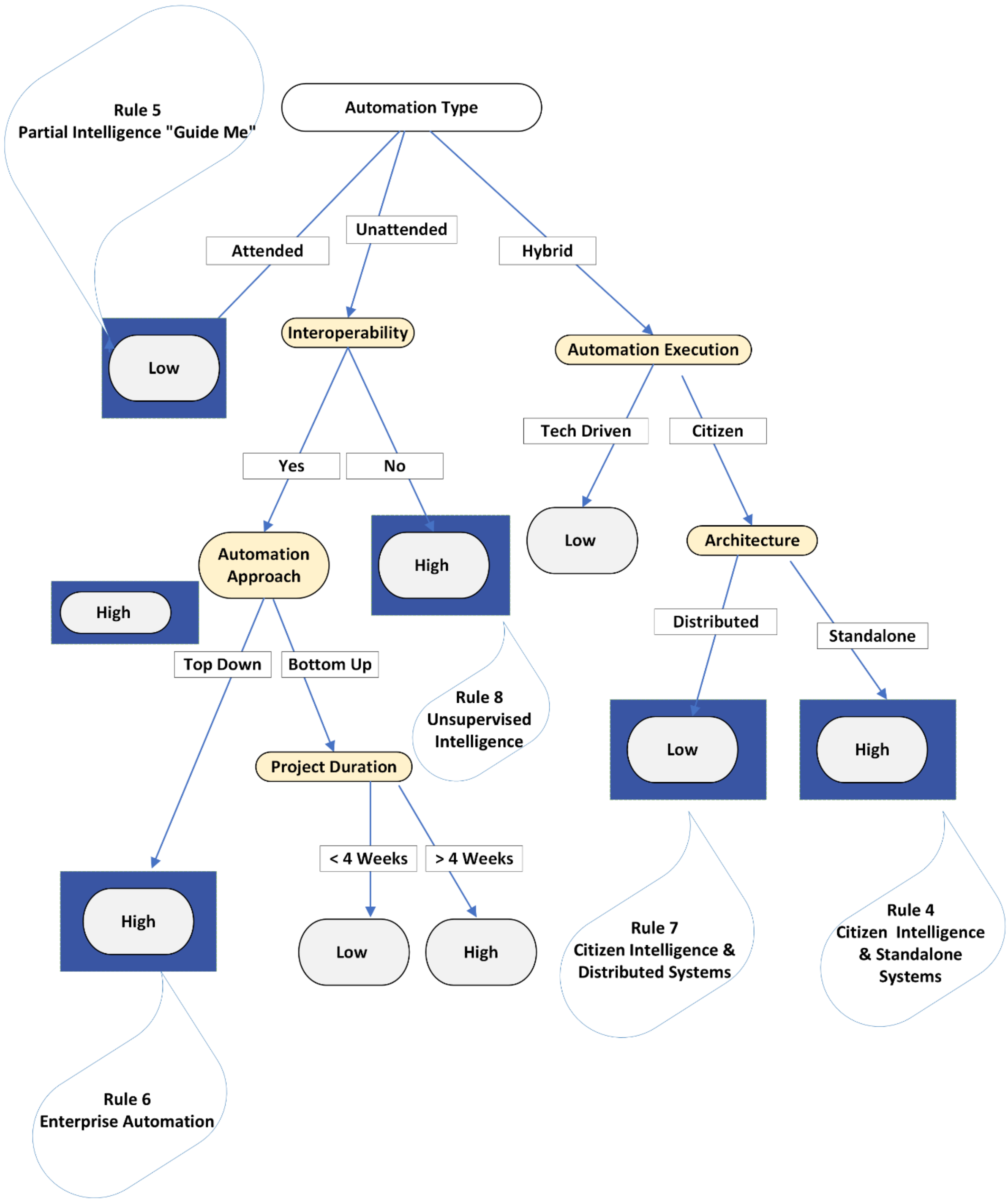

Figure 6: Tree 2: Decision Journey for Process Efficiency

The same data set of 176 intelligent process automation projects is used for computational experiments with the outcome variable "process efficiency" representing the success of intelligent process automation, as defined earlier for FTE reduction.

Prediction accuracy of the obtained tree from the training set is assessed by applying the rationale to predict the outcome "process efficiency" from unseen data in the testing set, which is a mutually disjoint set.

Tree 5 from Table 12 depicted in Figure 6 is the best representative tree; although prediction error is marginally higher than the least prediction error, it optimally satisfies the three key heuristics of low prediction error, parsimony, and consistency of the top-level variable.

From these top three variables, that is, "automation type," "interoperability," and "automation execution" impact **"process efficiency."** Higher the process efficiency, higher the success of intelligent process automation.

6.7.2 Robustness Check

I obtain the best representative tree i.e., tree 5 from table 1 through cross-validation using 10 folds, to ensure the robustness of the result I also perform a computational experiment on the best representative tree using percentage split test using 80%, which means that 80% of the 176 instances are trained and tested with reminder of 20% of instances. I find that both with 10-fold validation and 80% percentage split the best representative tree five yields same results.

6.8 Key Findings and Rules for Process Efficiency

The tree representation for "process efficiency" follows the same methodology as followed for the first outcome variable "FTE reduction" to yield context-specific rules that clearly articulate the emergent connections across levels to inform the IPA success factor of "process efficiency" and therefore constitute part of the multi-level theory. All the 10 information attributes characterizing initiatives in conjunction with the "process efficiency" decision are inputs to tree induction.

There are 5 rules derived in this study, as depicted in Figure 6: rules 1, 3, and 5 are the top rules that predict the factors of high process efficiency, resulting in successful IPA implementations, whereas rules 2 and 4 predict factors of low process efficiency, which leads to unsuccessful IPA implementations. Rules are presented in Table 12.

Table 13: Top context-specific rules discovered from decision tree induction (process efficiency)

No	Rule	Levels incorporated	Decision (process efficiency)	Instances classified
4	Citizen intelligence and standalone systems *Automation type ="hybrid" and Automation execution =" citizen" and IPA architecture = "standalone"*	*Governance Process Technology Complexity*	High	45
5	Partial intelligence "guide me" *Automation t= "attended"*	*Governance Process Technology Complexity*	Low	29
6	Enterprise automation *Automation type = "unattended" and Interoperability = "yes" and Automation approach = "top-down"*	*Governance Process Technology Complexity*	High	24
7	Citizen intelligence and distributed systems *Automation type = "hybrid" and Automation execution = "citizen "and IPA architecture = "distributed"*	*Governance Process Technology Complexity*	Low	22
8	Unsupervised intelligence *Automation type = "unattended" and Interoperability = "no"*	*Governance Process Technology Complexity*	High	22

6.8.1 Rule 4: Citizen Intelligence and Standalone Systems

Citizen automation happens when business users, without any knowledge or background in technical coding, develop IPA strategies to improve their work routines. In the current examination, citizen automators create solutions to replace human labor with machine intelligence, promoting digital transformation. Repetitive tasks that do not require intensive analysis are potential subjects of automation, i.e., require basic "if-then" thinking. As discussed earlier, when automation involves some user input, hybrid automation is recommended. From the 176 IPA implementations, when the outcome variable "process efficiency" is used and when the automation is hybrid and involves business user citizen automation with a standalone automation architecture, it generally leads to higher "process efficiency." In standalone automation, the business user is the owner of the process and completely understands the manual interventions required due to the subject matter expertise he/she has and can augment the machine effectively to improve the overall efficiency of the process. This was the main classification rule extracted from the tree as it classified most IPA implementations (45 IPA implementations). The general form of the rule is ***when the automation is hybrid and if automation is citizen development and the architecture is standalone, the process efficiency obtained in an IPA implementation is high.***

6.8.2 Rule 5: Partial Intelligence "Guide Me"

As explained earlier, attended automation generally occurs when there is no specific pattern identified for the business process and can only be performed by humans. From the studied 176 implementations for the outcome variable "process efficiency," it is observed that the process efficiency is very low when automation is attended. This was the main classification rule extracted from the tree as it classified most IPA implementations (29 IPA implementations). The general form of the rule is ***when automation is attended, then process efficiency obtained in an IPA implementation is low.***

6.8.3 Rule 6: Enterprise Automation

From my earlier examination it was been observed that unattended automation leads to greater FTE reduction and higher process efficiency; however, when the systems are more complicated and processes involve interacting and integrating with other processes and systems with greater interoperability, it was observed that the top-down approach to IPA is more suitable. The interoperability of systems and processes involve more sophisticated orchestration, scalability, and security of automation, indicating that the top-down approach is essential as it is implemented across the enterprise, and the company needs to follow a strategic top-down approach to automation as it is a key technology for any incumbent digitization journey that "pushes the envelope" on process redesign, enabling further process reengineering and thus requiring the top-down support of business owners at the C-level. From the studied 176 implementations for the outcome variable "process efficiency," it is observed that process efficiency is very high in unattended automation of interoperable business processes when the automation approach is strategic at the C-level, i.e., top-down. This was the main classification rule extracted from the tree as it classified most IPA implementations (72 IPA implementations). The general form of the rule is ***when the automation is unattended and if business processes are highly interoperable and if the automation approach is top-down, then the process efficiency obtained in an IPA implementation is high.***

6.8.4 Rule 7: Citizen Intelligence and Distributed Systems

Distributed automation systems require integration between multiple departments and processes, and the user input in the form of citizen development becomes overly complex. From the 176 IPA implementations, when the outcome variable is "process efficiency" and when the automation is hybrid and involves business user citizen automation with a distributed architecture, process efficiency is low. The manual interventions in a hybrid automation for distributed systems are difficult to augment machines and do not effectively increase the efficiency of the process. This was the main classification rule extracted from the tree as it classified most IPA implementations (24 IPA implementations). The general form of the rule is ***when the automation is hybrid and if automation is citizen development and the architecture is distributed, the process efficiency obtained in an IPA implementation is low.***

6.8.5 Rule 8: Unsupervised Intelligence

Unattended automation executes tasks of a process without any human involvement from the start to the end, where the tasks are mostly scheduled to start, or some event triggers the process to begin. Unattended automation manages manual tasks that **involve a specific pattern or specific set of steps that are meant to be followed**. From the studied 176 implementations for the outcome variable "process efficiency," it is observed that FTE reduction is very high when the automation is unattended, and the process is not interoperable. This was the main classification rule extracted from the tree as it classified most IPA implementations (80 IPA implementations). The general form of the rule is ***when automation is unattended, the process efficiency obtained in an IPA implementation is high.***

6.9 Examining Process Accuracy

Process accuracy is an important metric or outcome of intelligent process automation. A good process efficiency may be achieved by automating as many steps as possible in process automation; however, when an IPA robot is run, it must be highly accurate. Accuracy is defined as the ability to complete the process steps perfectly with zero errors (i.e., accuracy is 100%) (Gami et al., 2019).

The computational experiments are run again on the set of eleven predictors to generate decision rationale and theory development to derive most significant predictors impacting the process accuracy of intelligent process automation.

Table 14: Computational experiments to select best representative tree for "accuracy".

No.	Degree of pruning	Minimum instances at leaves	Communicability	Consistency	Accuracy
			Size of tree (# of leaves)	Top-most attribute	Prediction error (validation data)
1	High (0.25)	8	6	1. Top-down / bottom-up 2. Automation type	32.38%
2	Medium (0.5)	8	16	1. Top-down/ bottom-up 2. Automation type	30.11%
3	Low (0.75)	8	16	1. Top-down / bottom-up 2. Automation type	30.11%

4	High (0.25)	10	6	1. Top-down / bottom-up 2. Automation type	31.81%
5	Medium (0.5)	10	13	1. Top-down / bottom-up 2. Automation type	30.11%
6	Low (0.75)	10	13	1. Top-down / bottom-up 2. Automation type	28.40%
7	High (0.25)	12	6	1. Top-down / bottom-up 2. Automation type	31.81%
8	Medium (0.5)	12	9	1. Top-down / bottom-up 2. Automation type	30.11%
9	Low (0.75)	12	9	1. Top-down / bottom-up 2. Automation type	28.40%
10	High (0.25)	14	6	1. Top-down / bottom-up 2. Automation type	31.81%
11	Medium (0.5)	14	9	1. Top-down / bottom-up 2. Automation type	30.11%
12	Low (0.75)	14	9	1. Top-down / bottom-up 2. Automation type	28.40%

13	High (0.25)	16	6	1. Top-down / bottom-up 2. Automation type	31.81%
14	Medium (0.5)	16	9	1. Top-down / bottom-up 2. Automation type	30.11%
15	Low (0.75)	16	9	1. Top-down / bottom-up 2. Automation type	28.40%

6.9.1 Selecting the Best Representative Tree Process Accuracy

The following best representative tree provides a rationale indicating best predictors for success of IPA implementation in terms of process accuracy as an outcome.

The same data set of 176 intelligent process automation projects was used for computational experiments with an outcome variable "accuracy" representing success of intelligent process automation, as defined earlier for FTE reduction. Prediction accuracy of the obtained tree from the training set was assessed by applying the rationale to predict the outcome "accuracy" from unseen data in the testing set, which is a mutually disjoint set.

Tree 15, from Table 14 depicted in Figure 7, is the best representative tree that has lowest prediction accuracy; it optimally satisfies the three key heuristics of low prediction error, parsimony, and consistency of the top-level variable.

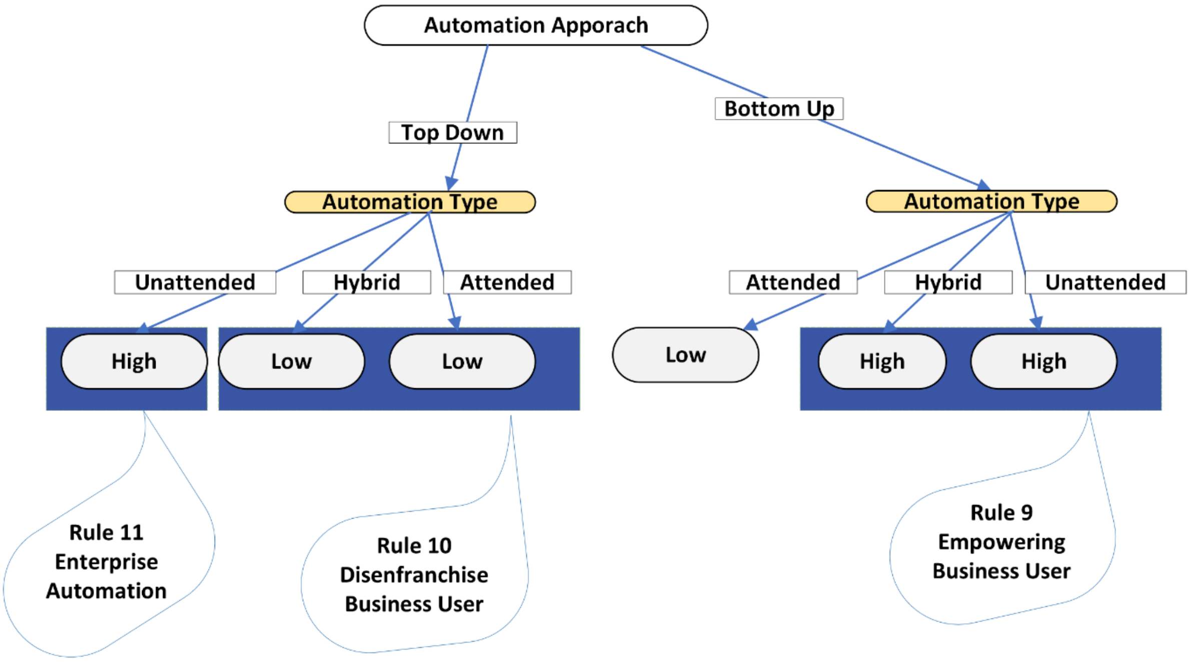

Figure 7: Tree 3: Decision Journey for Accuracy

From the top three variables, that is, "automation approach," "top-down/bottom-up," and "automation type" impact **"accuracy."** The higher the accuracy, the higher the success of intelligent process automation.

6.9.2 Robustness Check

The best representative tree, i.e., tree 15 from Table 14, was obtained **using 10-fold** validation**. To ensure the robustness of the result, a computational experiment was performed on the best representative tree using the percentage split test using 80%, which means that 80% of the 176 instances are trained and tested with reminder of 20% of instances. I find that both with 10-fold validation and 80% percentage split the best representative tree 15 yields same results.**

6.10 Key Findings and Rules for Process Accuracy

The tree representation for "accuracy" follows the same methodology as used for the first two outcome variables "FTE reduction" and "process efficiency" to yield context-specific rules that clearly articulate the emergent connections across levels to inform the IPA success factor of "accuracy" and therefore constitute part of the multi-level theory. All the 10 information attributes characterizing initiatives in conjunction with the "accuracy" decision are inputs to tree induction.

Three rules were derived in this examination, as depicted in Figure 7: rules 1 and 3 are the top rules that predict the factors for high accuracy, resulting in successful IPA implementations, and rule 2 is the top rule that predicts the factor for low accuracy, resulting in lower success of IPA implementation. Rules are presented in Table 14.

Table 15: Top context-specific rules discovered from decision tree induction (accuracy)

No	Rule	Levels incorporated	Decision (accuracy)	Instances classified
9	Empowering business user *Automation approach = "bottom-up"* *Automation type = "hybrid" or "unattended"*	*Governance* *Process* *Technology* *Complexity*	High	93
10	Disenfranchise business user *Automation approach = "top-down" and* *Automation type = "attended" or "hybrid"*	*Governance* *Process* *Technology* *Complexity*	Low	41
11	Enterprise automation *Automation approach = "top-down" and* *Automation type = "unattended"*	*Governance* *Process* *Technology* *Complexity*	High	25

6.10.1 Rule 9: Empowering Business User

As discussed earlier, bottom-up automation is driven by people at the business process level (business owner) who are empowered to define automation, with an all-inclusive approach that results in successful automation of a process; this kind of bottom-up approach will also make FTEs more productive. From the 176 IPA implementations and when "accuracy" is considered as the outcome variable, due to the inclusive approach, the business owners are in full control of the process steps, and they either define all the steps of unattended automation or they augment the machine effectively with the steps that cannot be automated, thereby achieving high accuracy. This was the main classification rule extracted from the tree as it classified most IPA implementations (93 IPA implementations). The general form of the rule is ***when automation approach is bottom-up and automation type is either unattended or hybrid, the accuracy of IPA implementation obtained is higher.***

6.10.2 Rule 10: Disenfranchise Business Owner

It is observed that when the automation approach is top-down, i.e., top management or a C-level takes a decision to automate the process, and the automation type is either attended or hybrid, i.e., involves human intervention, and the business user is not adequately involved, and the accuracy takes a hit. From the studied 176 implementations for the outcome variable "accuracy," it is observed that accuracy is very low when the automation is top-down and involves the business user intervention without empowering them to take decisions. This was the main classification rule extracted from the tree as it classified most IPA implementations (41 IPA implementations). The general form of the rule is ***when automation approach is top-down and the automation type is either attended or hybrid, the accuracy of IPA implementation obtained is lower.***

6.10.3 Rule 11: Enterprise Automation

As explained earlier, when decisions are made top-down, there is an alignment between departments within the enterprise toward the automation goal, and especially when the process across the enterprise is understood well and automated in such a way that there is no human intervention, the accuracy of the IPA implementation is high. This was the main classification rule extracted from the tree as it classified most IPA implementations (25 IPA implementations). The general form of the rule is ***when the automation approach is top-down and the automation type is unattended, then the accuracy of IPA implementation obtained is higher.***

6.11 Conclusion

This chapter details the process of decision tree induction and how it is useful in deriving insights and patterns and explains data partitioning to identify combinations of information attributes associated with similar outcomes using partitioning and data validation using the C4.5 algorithm using the WEKA machine learning platform. It also explains the need for computational experiments to prune decision trees with varying confidence levels and number of tree instances, which helps in the reduction of the tree size by reducing the part of tree that has little power to classify data and reducing the noise or overfitting. It also defines the three key heuristics (high predictive accuracy, communicability or high parsimony, and high reliability), which helps identify the best representative tree. As a next step, the analysis for three outcomes of abduction in stage 1 of the research methodology, namely, FTE reduction, process efficiency, and process accuracy, is explained through induction of decision trees, selecting the best representative tree, and the rules obtained. Then, eleven rules are derived as a result of decision tree induction.

The next chapter will explain the process of deriving one composite index IPA success through a principal component analysis formative construct and by performing decision tree induction to obtain the best representative tree for IPA success and the rules derived.

7 COMPOSITE MEASURE OF IPA SUCCESS

In this chapter, I explain the process of obtaining a composite measure of IPA Success derived from the three measures of outcome obtained in the previous chapter. This measure, which is derived from principal component analysis (PCA), allows for a more holistic view of success, and can provide a more accurate representation of the critical success factors of intelligent process automation (IPA) implementation. The measure is also utilized for the decision tree induction to frame rules for critical success factors of IPA implementation.

7.1 Formative vs Reflective Constructs

From the earlier examination outlined in Chapter 6, I observed that I have three different measures of outcome, that is, Full-time Equivalent (FTE) Reduction, Process Efficiency, and Accuracy. However, the objective of this research paper was both to study the individual outcomes of IPA Success and to measure a single-index measure called IPA Success. To achieve this single-index IPA Success, one among formative or reflective construct as a latent variable was taken into account. From the literature review, it was observed that most of the researchers lean toward focusing more on the structural model rather than measurement models; that is, by fully considering the relationship between measures and their latent constructs (Jarvis et al., 2003), such errors in measurement models will lead to measurement errors and in turn effect the structural model (MacKenzie et al., 2005). However, some researchers considered all constructs alike, regardless of whether the construct is reflective or formative (Chin, 1998), and such misspecifications of constructs as formative or reflective will lead to Type 1 and Type 2 errors. Now, let us look at what these reflective and formative constructs are and the relevance to this study of identifying success factors of IPA implementation and defining a success indication of IP Success.

7.1.1 Reflective Construct

According to (Jarvis et al., 2003), a reflective construct is the one where changes in underlying latent constructs are hypothesized to cause changes in the indicative measures, that is, when measures are used to examine the underlying construct that is not observable and is referred to as reflective indicator or effect indicator (Edwards and Bagozzi, 2000).

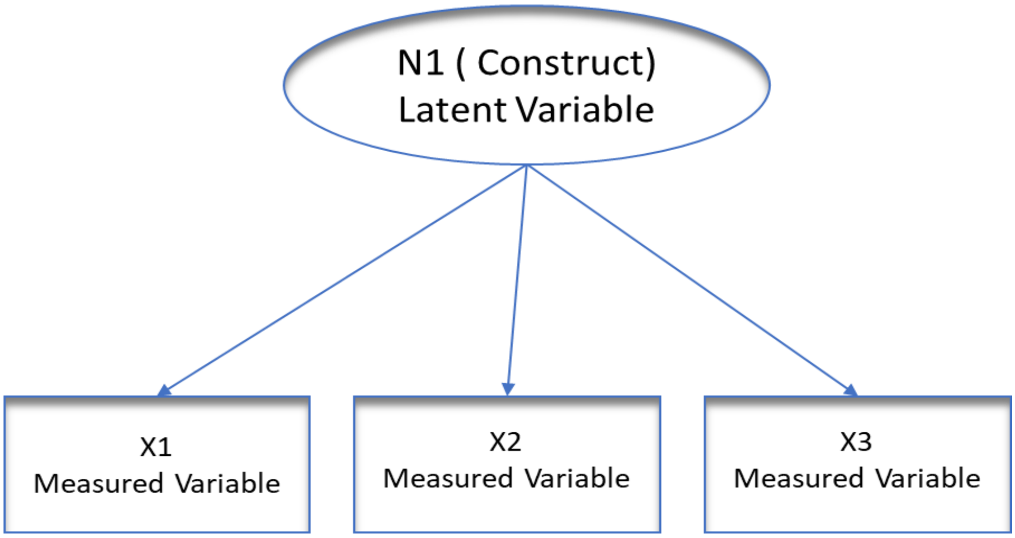

Figure 8: Representation of reflective construct

7.1.2 Formative Construct

In simple terms, the formative constructs are a composite of multiple measures (MacCallum and Browne, 1993), which means variations in formative construct effect changes in the underlying constructs. My research paper focuses on the success factors of IPA through a single measure or index known as IPA Success, which is operationalized by three measures: FTE

Reduction, Process Efficiency, and Accuracy. Each of them captures differing aspects of IPA Success, which results in *formative construct (Petter et al., 2007).*

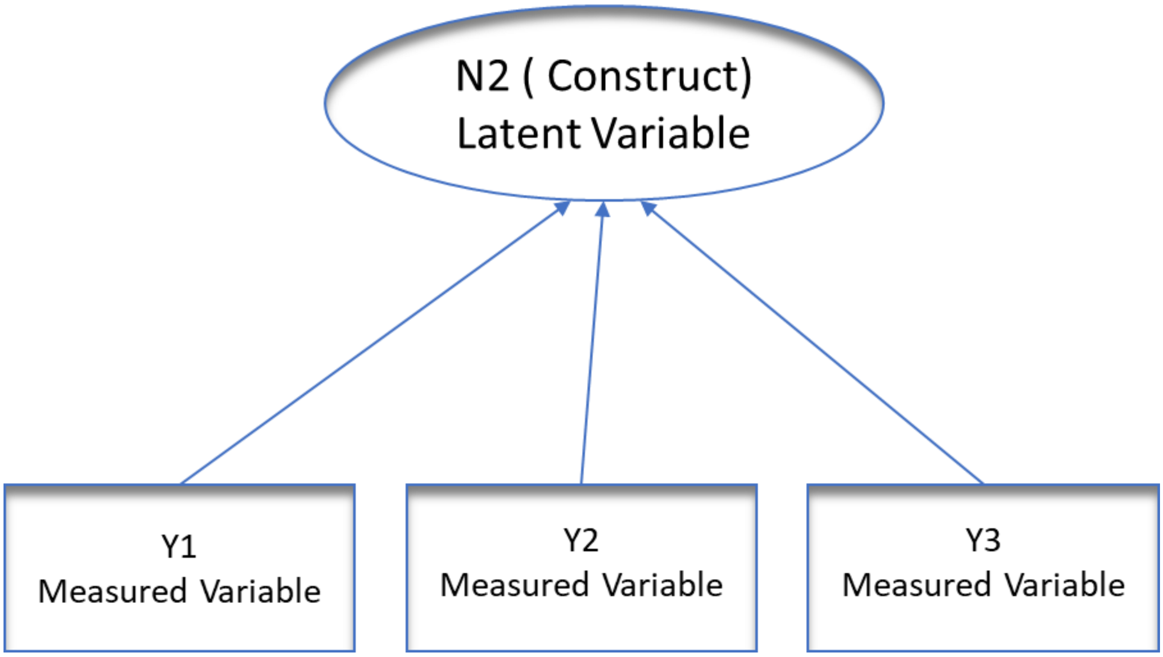

Figure 9: Representation of formative construct

I modeled my focal variable as formative construct since it meets the criteria of coverage of construct domain and lack of covariance among indicators (Diamantopoulos and Winklhofer, 2001). First, each item makes unique contributions to the constructs and can be viewed as "forming" them: for example, gross domestic product, which measures country's economic performance considering the value of all goods and services produced within its borders. Second, an increase in any one item does not necessarily increase others. For example, an increase in sales growth does not necessarily imply an increase in profitability. Finally, items comprising each construct are distinct and not interchangeable.

A more comprehensive measure of the construct to identify the critical success factors of IPA is postulated in this study, which is very context-specific deriving from multiple dimensions represented by three measures of outcomes, and formative construct is the best way to achieve the comprehensive measure “IPA Success.”

7.2 Principal Component Analysis

Formative construct retains the unique variance in each measure as against the three outcome measures discussed in Chapter 6; hence, I use PCA to derive the IPA Success Index. This method helps reduce the dimensionality of the measures (Chin, 1995).

PCA is a statistical technique used to identify patterns in data. It is a dimensionally reduced method that is used to transform many variables into a smaller number of uncorrelated variables, known as principal components.

PCA works by identifying the direction of maximum variance in the data and then projecting the data onto this direction, creating a new variable (i.e., principal component) that captures as much of the variance in the original data as possible. This process is then repeated for the remaining directions of maximum variance, creating additional principal components.

The resulting principal components are linear combinations of the original variables and are orthogonal to each other (i.e., uncorrelated). The first principal component captures the most variation in the data, with each subsequent component capturing progressively less. The number of principal components retained is determined by the amount of variance that needs to be explained, as well as the interpretability of the components.

7.3 Robustness Check

The values obtained from the IPA Success Index using PCA were compared with those in structural equation modelling (SEM) using smart PLS, and the values with both methods are identical; hence, the robustness of the result is established.

7.4 Data Analysis for Outcome Predictor IPA Success

IPA Success is one of the key success indicators of IPA, and it is formative construct derived from PCA of three outcome variables discussed in Chapter 6, namely, FTE Reduction, Process Efficiency, and Accuracy. The IPA Success will act as one single index for the success of IPA. The values of the IPA Success Index obtained through PCA is divided into three sets in ascending order of success index, with first being 34% (high), next 33% (medium), and remaining 33% (low).

The data set for IPA Success undergoes the same process of computational experiments, pruning through 10-fold validation with the same set of 16 predictors (**governance level, process level, technology level, and complexity level)** as discussed earlier in Chapter 6 to identify best representative tree and generate decision rationale, thereby aiding the development of theory to identify critical factors effecting IPA Success. Computational experiments for IPA Success are depicted in **Table 15.**

Table 16: Computational experiments to select best representative tree for "IPA Success."

Mechanisms to Detect Overfitting			**Three Key Heuristics to Choose the Best Representative Model**		
No.			COMMUNICABILITY	CONSISTENCY	ACCURACY

	Degree of pruning	Minimum instances at leaves	Size of tree (# of leaves)	Topmost Attribute	Prediction Error (Validation Data)
1	High (0.25)	8	4	1. Automation Type	38.00%
2	Medium (0.5)	8	9	1. Automation Type 2. Complexity	38.63%
3	Low (0.75)	8	12	1. Automation Type 2. Complexity	39.7%
4	High (0.25)	10	4	1. Automation Type	39.70%
5	Medium (0.5)	10	7	1. Automation Type 2. Complexity	41.40%
6	Low (0.75)	10	10	1. Automation Type 2. Complexity	41%
7	High (0.25)	12	4	1. Automation Type	25%
8	Medium (0.5)	12	7	1. Automation Type 2. Complexity	42%
9	Low (0.75)	12	10	1. Automation Type 2. Complexity	41.4%
10	High (0.25)	14	4	1. Automation Type	40%
11	Medium (0.5)	14	7	1. Automation Type 2. Complexity	42%
12	Low (0.75)	14	10	1. Automation Type 2. Complexity	41.4%
13	High (0.25)	16	4	1. Automation Type	40.3%
14		16	7		42%

	Medium (0.5)			1. Automation Type 2. Complexity	
15	Low (0.75)	16	7	1. Automation Type 2. Complexity	41.4%

7.4.1 Selecting the Best Representative Tree "IPA Success"

The best representative tree shown in Figure 10 provides a rationale indicating best predictors for the success of IPA implementation in terms of overall IPA Success as an outcome.

The same data set of 176 IPA projects is used for computational experiments, this time with an outcome variable "IPA Success" representing the success of intelligent process automation.

Prediction accuracy of obtained tree from training set is assessed by applying the rationale to predict the outcome "IPA Success," from unseen data in the testing set, which is mutually disjoint set.

Table 16Tree 3 from Table 16 depicted in Figure 10 is the best representative tree, which optimally satisfies the three key heuristics of low prediction error, parsimony, and consistency of top-level variable.

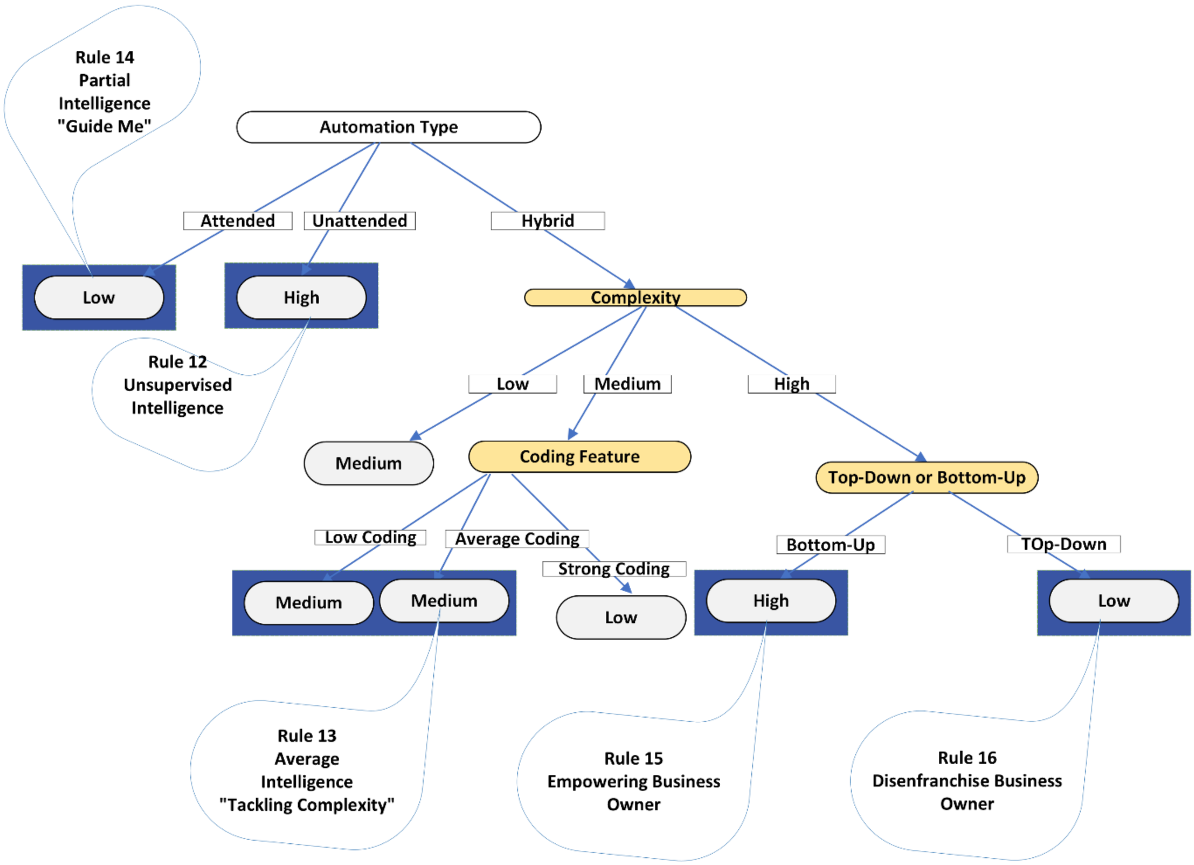

Figure 10: Tree Four: Decision Journey for IPA Success

The higher the impact of "IPA Success", as shown in the top variables "Automation Type" and "Complexity" in Figure 10, the higher the success of IPA.

7.4.2 Robustness Check

The best representative tree, that is, Tree 3 from

Table 16 was obtained through cross-validation using 10-folds, to ensure the robustness of the result. A computational experiment on the best representative tree was also performed using the percentage split test using 80%, which means that 80% of the 176 instances are trained and tested with reminder 20% of instances. I find that both with 10-fold validation and 80% split the best representative Tree 5 yields the same results.

7.5 Key Findings & Rules for IPA Success

The tree representation for "IPA Success" follows the same methodology as I have discussed in Chapter 6 to yield context-specific rules that clearly articulate the emergent connections across levels to inform the IPA Success Factor "IPA Success" question and therefore constitute part of my multilevel theory. All the ten information attributes characterizing initiatives in conjunction with the "IPA Success" decision are inputs to tree induction.

There are five rules derived in this study as depicted in Figure 3: Rules 1 and 4 are the rules that predict the factors for high "IPA Success" resulting in successful IPA implementations, Rule 2 predicts the factors of medium "IPA Success", and rules 2 and 5 are the rules that predict the factors for low "IPA Success." Rules are presented in Table 16.

Table 17: Top context-specific rules discovered from decision tree induction "IPA Success."

No.	Rule	Levels Incorporated	Decision (IPA Success)	Instances Classified
12	Unsupervised Intelligence *Automation Type ="Unattended"*	*Governance* *Process* *Technology* *Complexity*	High	70
13	Average Intelligence "Tackling Complexity" *Automation Type ="hybrid" and Complexity = "Medium" and Coding Feature ="low Coding" or "Average Coding"*	*Governance* *Process* *Technology* *Complexity*	Medium	41
14	Partial Intelligence "Guide Me" *Automation Type ="Attended"*	*Governance* *Process* *Technology* *Complexity*	Low	29
15	Empowering Business Owner *Automation Type ="hybrid" and Complexity = "High" and*	*Governance* *Process* *Technology*	High	16

	Automation Approach = "Bottom-Up"	*Complexity*		
16	Disenfranchise Business Owner *Automation Type ="hybrid" and* *Complexity = "High" and* *Automation Approach = "Top-Down"*	*Governance* *Process* *Technology* *Complexity*	Low	12

7.5.1 **Rule 12—Unsupervised Intelligence**

Unattended automation executes tasks of a process without any human involvement from start to end. They are mostly scheduled to start the process or when there is some event that triggers the process to begin. They run in the background and pass their product to either humans or some other machine. Unattended automation manages manual tasks that **involves specific pattern or specific set of steps that are meant to be followed**. From the studied 176 implementations for the outcome variable "IPA Success," it is observed that the IPA Success is very high when the automation is unattended. This was the main classification rule extracted from the tree as it is classified in most IPA implementations (70 IPA implementations). The general form of the rule is ***when automation is unattended, then the IPA Success obtained is high.***

7.5.2 **Rule 13—Average Intelligence "Tackling Complexity"**

This rule specifically illustrates that when the automation is in hybrid mode, humans augment machines to execute the process. When the coding is either average or low, the extent of automation is moderate or medium for the 176 IPA implementations studied. This was the main classification rule extracted from the tree as it is classified in most IPA implementations (41 IPA implementations). The general form of the rule is even ***when the coding is either low or average***

and the business process is of moderate complexity and still needs humans to augment automation, then the IPA Success obtained is moderate.

7.5.3 **Rule 14—Partial Intelligence "Guide Me"**

Attended automation generally collaborates with humans, handling certain tasks within longer, more complex workloads or processes that cannot be fully automated from start to finish. Attended automation occurs when there is no specific pattern identified for business process and can only be performed by humans. From the studied 176 implementations for the outcome variable "IPA Success," it is observed that the IPA Success is very low when the automation is attended. This was the main classification rule extracted from the tree as it is classified most IPA implementations (29 IPA implementations). The general form of the rule is ***when automation attended, the FTE Reduction obtained in an IPA implementation is low.***

7.5.4 **Rule 15—Empowering Business Owner**

When automation involves some user input, hybrid automation is recommended. I also understand bottom-up automation as something that is driven by people at the business process level who are empowered to give ideas as they have the complete knowledge of domain and the gaps to be automated, with an all-inclusive approach that results in successful automation of a process. In the context of this specific rule, hybrid automation involves humans or business owners who drive the automation through bottom-up approach; specifically in the complex processes, it can bridge the gaps where machine cannot automate and hence lead to high IPA Success.

From the 176 IPA implementations analyzed in this study, I considered outcome variable as "IPA Success" and found that when the automation type is hybrid, that is, combination of attended and unattended, when the **business process owners are involved in decision making of automation,** and when process complexity is high, the inclusive bottom-up approach IPA Success is high. This was the main classification rule extracted from the tree as it is classified in most IPA implementations (16 IPA implementations). The general form of the rule is ***the IPA Success obtained is high when automation is hybrid and when dealing with complex process, and if driven by bottom-up approach.***

7.5.5 **Rule 16—Disenfranchise Business Owner**

As in the case of rule 4, hybrid automation takes place when the process needs to be augmented with human intelligence, where specific inputs from human are required to make the automation of the process successful. In this specific scenario of Rule 5 where the complexity of the process is high, it requires a human intervention. These human interventions should be handled by business users with complete domain knowledge of the process being automated to fill the gaps that machine is not able to achieve. In the scenario of complex process, the human intervention is through top-down approach, which means that when someone who is not familiar with the process (not a business process owner) and is positioned higher in the hierarchy is involved, then the success of IPA is low.

From the 176 IPA implementations analyzed in this study, I considered outcome variable as "IPA Success" and found that when the automation type is hybrid, that is, combination of attended and unattended, when the **business process owners are not involved in decision making of automation,** and when process complexity is high, the non-inclusive bottom-up approach IPA Success is low. This was the main classification rule extracted from the tree as it is classified in most IPA implementations (12 IPA implementations). The general form of the rule is ***the IPA Success obtained is high when automation is hybrid and when dealing with complex process, and if driven by top-down approach.***

7.6 Conclusion

In this chapter thus far, I explain the differences between the formative and reflective constructs and a rationale behind using constructive construct using the PCA. Through PCA, I obtain the comprehensive construct that measures the IPA Success for the sample of 176 live IPA implementations. Once the single-outcome measure IPA Success is obtained, the same decision tree induction process is performed through the analysis of the data and a best representative tree is obtained for the composite measure. From the best representative tree, I derive the critical success factors and five context-specific rules impacting the success or failure of the IPA implementation. In the next chapter, I develop the proposition by comparing and contrasting 16 rules derived so far.

8 ABDUCTING AWAY TO DEVELOP INSIGHTS AND PROPOSITIONS

In this chapter, I first compare and contrast the rules derived from decision trees for IPA Success, then I represent the rules separately for high and low IPA Success. I then abduct away from these rules to form the propositions.

8.1 Compare and Contrast Rules from Decision Trees

In Chapter 6, through the decision tree induction, I examine the three measures of outcome (FTE Reduction, Process Efficiency, and Process Accuracy), and in Chapter 7, I obtain a single measure of outcome IPA Success, then I compare and contrast the rules obtained from the examination of all four outcomes as part analysis one in Chapter 6 and analysis two in Chapter 7 to identify the insights and propositions for my research questions on the critical success factors of Intelligent Process automation (IPA) Success.

Table 18: Compare and Contrast Rules Across IPA Outcomes

	Analysis 1			**Analysis 2**
Automation Success	FTE Reduction	Process Efficiency	Accuracy	IPA Success
High	Unsupervised Intelligence Automation Type = “Unattended”	Unsupervised Intelligence Automation Type = “Unattended”		Unsupervised Intelligence Automation Type = “Unattended”
	Empowering Business User Automation Type = “Hybrid” and Automation Approach = “Bottom-up” and Complexity = “High” or “Medium” or “Low”		Empowering Business User Automation Approach = “Bottom-Up” Automation Type = “Hybrid” or “Unattended”	Empowering Business Owner Automation Type = “hybrid” and Complexity = “High” and Automation Approach = “Bottom-Up”
		Citizen Intelligence and		

		Standalone Systems Automation Type =“Hybrid” and Automation Execution = “Citizen “and IPA Architecture = “Stand Alone”		
		Enterprise Automation Automation Type = “Unattended” and Interoperability = “Yes” and Automation Approach = “Top-Down”	Enterprise Automation Automation Approach = “Top-Down” and Automation Type = “Unattended”	
Low	Partial Intelligence “Guide Me” Automation Type = “Attended”	Partial Intelligence “Guide Me”" Automation Type = “Attended”		Partial Intelligence “Guide Me” Automation Type = “Attended”
			Disenfranchise Business User Automation Approach = “Top-Down” and Automation Type = “Attended” or “Hybrid”	Disenfranchise Business Owner Automation Type = “hybrid” and Complexity = “High” and Automation Approach = “Top-Down”
		Citizen Intelligence & Distributed Systems Automation Type = “Hybrid” and Automation Execution = “Citizen “and IPA Architecture = “Distributed”		
Medium				Average Intelligence "Tackling Complexity" Automation Type = “Hybrid” and Complexity = “Medium” and Coding Feature = “Low Coding” or “Average Coding”

8.2 Insights

From the context-specific rules induced through decision trees, I now abduct away to present the insights based on the compare-and-contrast table represented in Table 18. These insights form the basis for making insightful decisions for IPA practitioners. Following are the insights on critical success factors impacting the IPA implementation.

First, I considered the entire repertoire of context-specific rules and the theoretical levels and predictors they encompass. As seen in Table 11, Table 13, Table 15, and Table 17, for all four outcomes of IPA Success, I observed that there are some top-ranked predictors that results in high IPA implementation success; abstracting away from this observation, I present the first insight as shown below.

Table 19 : Observation and Insight 1.

Observation	**Unattended automation type design is a predictor for the top-ranked rule that results in high IPA implementation success in all four outcomes (i.e., FTE Reduction, Process Efficiency, Process Accuracy, and IPA Success)**
Insight	S1: Unattended automation-type design is a **necessary** predictor for the top-ranked rule that results in high IPA implementation success for all outcomes.

Second, I observe that there are some top-ranked predictors that result in low IPA implementation success; abstracting away from this observation, I present the second insight as shown below.

Table 20: Observation and Insight 2.

Observation	**Attended automation-type design is a predictor for the top-ranked rule that results in low IPA implementation success in all four outcomes**

	(i.e., FTE Reduction, Process Efficiency, Process Accuracy, and IPA Success)
Insight	S2: Attended automation-type design is a **necessary** predictor for the top-ranked rule that results in low IPA implementation success for all outcomes.

Third, I observe that there is a second-ranked predictor that results in high IPA implementation success for three out of four outcomes; abstracting away from this observation, I present the third insight as shown below.

Table 21: Observation and Insight 3.

Observation	**Bottom-Up automation approach is a predictor for the second-ranked rule that results in high IPA implementation success for three outcomes (i.e., FTE Reduction, Process Accuracy, and IPA Success)**
Insight	S3: Bottom-Up automation approach is a **necessary** predictor for the second-ranked rule that results in high IPA implementation success for all outcomes except Process Efficiency.

Next, I observe that a specific predictor Top-Down, with a specific combination resulting in high IPA implementation success for two out of four outcomes; abstracting away from this observation, I present the fourth insight as shown below.

Table 22: Observation and Insight 4.

Observation	**Top-Down automation approach is a predictor that results in high IPA implementation success for two out of the four outcomes (i.e., Process Accuracy and Process Efficiency), when combined with the interoperable processes and Unattended automation type**
Insight	S4: For interoperable and unattended processes, Top-Down automation approach is an **important** predictor of high IPA implementation success for process accuracy and efficiency.

Next, I observe that some predictor combinations impact high IPA implementation success for a single outcome and are not present in any of the other outcomes. Citizen execution in standalone architecture results in high IPA implementation success for the outcome process efficiency; abstracting away from this observation, I present the fifth insight as shown below.

Table 23: Observation and Insight 5.

Observation	**Citizen automation execution is a predictor that results in high process efficiency IPA outcome, when the technology architecture is standalone and automation type is Hybrid. Citizen automation predictor is not present in other three IPA implementation success outcomes (i.e., FTE Reduction, Process Accuracy, and IPA Success)**
Insight	S5: Citizen automation execution is an **important** predictor that results in high process efficiency when combined with standalone architecture and Hybrid automation; it is not present in any of the other IPA implementation success outcomes.

Next, I observe that some predictor combinations impact low IPA implementation success for a single outcome and is not present in any of the other outcomes. Citizen execution in distributed architecture results in high IPA implementation success for the outcome process efficiency; abstracting away from this observation, I present the sixth insight as shown below.

Table 24: Observation and Insight 6.

Observation	**Citizen automation execution is an important predictor that results in low process efficiency IPA outcome; when the technology architecture is distributed, and automation type is Hybrid. Citizen automation predictor is not present in other three IPA Implementation success outcomes (i.e., FTE Reduction, Process Accuracy, and IPA success)**
Insight	S6: Citizen automation execution is an **important** predictor that results in low process efficiency when combined with distributed architecture and Hybrid automation; it is not present in any other IPA implementation success outcomes.

Through the process of abducting away, I arrive at insights that offer immense value to the practitioners implementing IPA in the organizations, and I look at the rules derived from decision trees and their combinations and impact across all the outcomes of IPA implementation to make these recommendations on the insights.

8.3 Propositions

Building on the context-specific rules induced through decision tree induction, I now move forward to generic explanations facilitated by abduction. For example, when discussing the hiring of employees in firms, (Pentland, 1999) argued that an observation enables storytelling from the point of view of a specific stakeholder. Abducting away from a specific new employee's perspective, a fabula can serve as the basis for generic propositions which reveal the underlying structure to a set of events and their interrelationships in terms of who did what and how people in general are hired.

Along similar lines, in this study, I make observations relying on the context-specific rules extracted from the best representative tree and the insights derived so far. In this phase of theory development, I abduct away from these specific observations to uncover the underlying structure of interrelationships of predictors across levels.

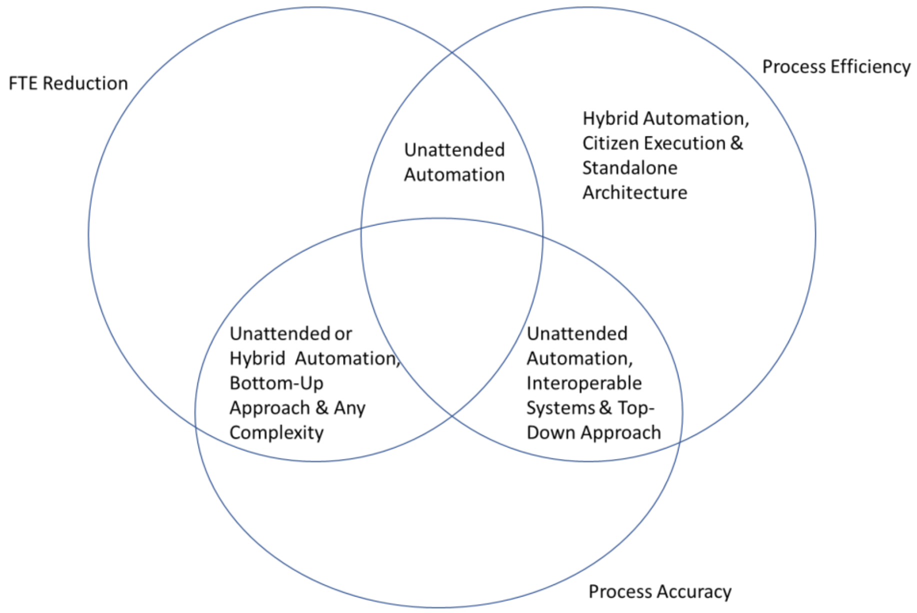

Figure 11: Compare and Contrast Rules for High IPA Success

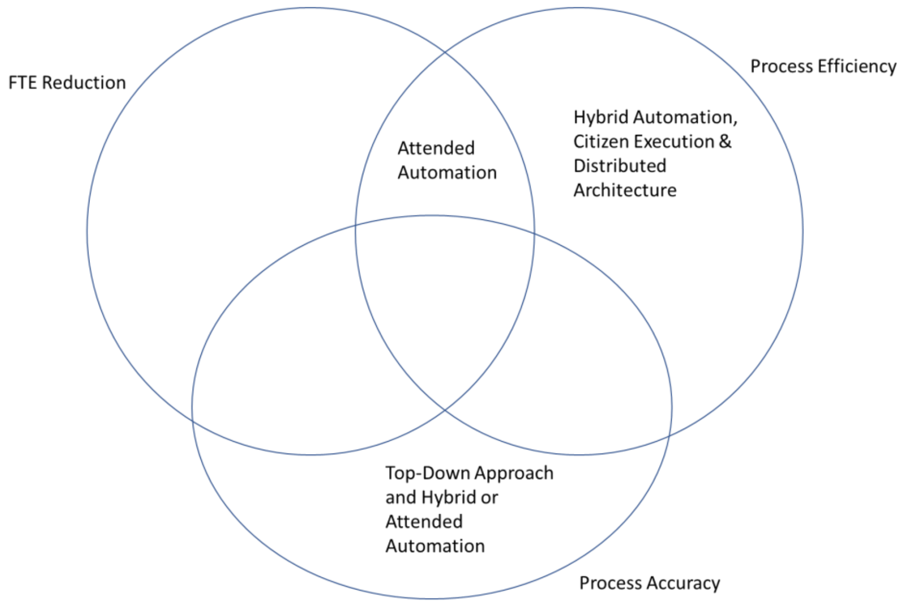

Figure 12: Compare and Contrast Rules for Low IPA Success

We articulate the individual rules or critical success factors impacting the high or low success of one or combination of IPA outcomes (FTE Reduction, Process Efficiency, Process Accuracy, and Overall IPA Success) as presented in Table 18, and Figure 12 and ,Figure 11. I articulate six generic propositions for critical success factors impacting the IPA Success, which broaden the generalizability of our research.

First, I considered the entire repertoire of context-specific rules, Insights, the theoretical levels, and predictors they encompass. As seen in Table 11, Table 13, Table 15, and Table 17, for all four outcomes of IPA Success, I observe several instances of rules with predictors across all theoretical levels. Abstracting away from this observation, I offer our first proposition as shown in Table 25.

Table 25: Observation and Theoretical Proposition 1.

Observation	**Predictors informing success of IPA implementations lie at Governance, Process, Technology, and Complexity levels**

Proposition	P1: Predictors informing success of IPA Implementation are present across multiple levels of analysis

Second, I compared and contrasted context-specific rules that collectively explain the predictors or critical factors impacting high and low IPA Success in Figure 11 and Figure 12. I observed several rules explaining critical factors of high and low IPA success with various predictors. Abducting away from these two observations, I develop our second proposition as shown in Table 26

Table 26: Observation and Theoretical Proposition 2.

Observation	**Predictors informing high success of IPA implementation are Unattended Automation, Bottom-Up Approach, Interoperable Systems, Citizen Execution, and Standalone Architecture.** **Predictors informing low success of IPA implementation are Attended Automation, Top-Down Approach, and Distributed Architecture**
Proposition	P2: The combination of predictors informing high success of IPA implementation is different from that of predictors informing low success of IPA implementation

Third, I scrutinized the interrelationships among predictors across various levels. I observed that only certain predictors are relevant for explaining IPA success based on the values of other predictors. Abducting away from these observations, I develop our third proposition in Table 27.

Table 27: Observation and Theoretical Proposition 3

Observation	**Citizen Automation Execution (Governance Level) is a significant predictor informing high success of IPA implementation only when the technology architecture (Technology Level) is Standalone and automation type (Complexity Level) is Hybrid.** **Citizen Automation Execution (Governance Level) is a significant predictor informing Low Success of IPA Implementation only when the**

	technology architecture (Technology Level) is Distributed and Automation Type (Complexity Level) is Hybrid.
Proposition	P3: For predictors informing high or low success of IPA implementations, the predictors at one level influence the inclusion or exclusion of predictors at other levels

Next, considering multiple contingencies, a few predictors could have a strong, dominating influence and almost operate in isolation. Alternatively, multiple contingencies could also combine to have a reinforcing influence on outcomes. I made two observations along these lines. Abducting away from these observations, I uncover two types of underlying interrelationships between predictors at multiple levels and I develop the fourth proposition in Table 28.

Table 28: Observation and Theoretical Proposition 4

Observation	**For success in IPA Implementation, Automation Type is a dominant predictor.** **For success in IPA Implementation, Automation Approach, Automation Execution, Process Complexity, IPA Architecture are reinforcing predictors**
Proposition	P4: Multiple predictors informing the success of IPA implementation may interact in such a way that either a few predictors dominate or reinforce in their combined influence on IPA Success.

Next, I have examined four measures of the IPA Success (FTE Reduction, Process Efficiency, Process Accuracy, and Overall IPA Success), and I scrutinize the one or combination of predictors on overall effect on all the four measures of IPA Success. I see that one of the most dominating combinations of predictors of the fourth and significant measure (Overall IPA Success) may or may not impact the other three outcome measures of IPA Success. Abducting away from these observations, I develop our fifth proposition in Table 29.

Table 29: Observation and Theoretical Proposition 5

Observation	**Unattended automation is the single dominating predictor informing the outcome measures Overall IPA Success, FTE Reduction, and Process Efficiency except Process Accuracy.** **The combination predictors of Hybrid automation and Bottom-Up automation approach informing the Overall IPA Success, FTE Reduction, and Process Accuracy except Process Efficiency**
Proposition	P5: One or combination of predictors informing Overall IPA Success may or may not impact other specific outcomes of IPA success

Finally, I see that some predictors informing the successful IPA implementation outcome, for example, Bottom-Up automation approach, the opposite, that is, Top-Down automation approach, do not necessarily imply that they will lead to failure of IPA implementation outcome; abducting away from these observations, I develop the fifth proposition shown in Table 29.

Table 30: Observation and Theoretical Proposition 6

Observation	Bottom-Up Automation approach leads to successful IPA implementation outcome when Automation Type is Hybrid and complexity is high; this does not mean that Top-Down leads to Unsuccessful IPA implementation outcome. **Top-Down Automation approach leads to successful IPA implementation outcome, when the business processes are Interoperable and Automation Type is Unattended**
Proposition	P6: If the presence of a predictor is necessary for successful IPA implementation outcome, the opposite does not imply that I will lead to failure.

In summary, by making systematic observations, our abductive approach facilitates the journey from context-specific rules to generic explanations in the form of propositions. In doing so, I build on prior exemplars of theory building with multiple contingencies and leverage multi-level theorizing to shed new light on critical success factors of IPA implementation.

Thus far in this chapter, I focused on abducting away from the rules derived in Chapters 6 and 7 to develop six insights and six meta propositions for the success of IPA implementations. In the next chapter, I validate the rules and propositions through econometrics establishing the extent of effect of each predictor and causality.

9 VALIDATION ANALYSIS THROUGH ECONOMETRIC METHODOLOGY

9.1 Introduction

In the previous chapters, we have discussed how a configurational view of causality has been established using decision tree induction. The configurational view of causality refers to a perspective that takes into account multiple configurations or combinations of causal factors that can lead to an outcome or result. It relies on the notion of complex causality using conjunctural, equifinal, and asymmetric relationships (El Sawy et al., 2010).

Decision trees are useful tools for representing and analyzing these configurations. They are graphical representations of a set of decisions and their possible consequences (equifinality). They consist of nodes that represent decision points, branches that represent possible choices, and leaves that represent outcomes. For example, the automation type node for successful implementation of intelligent process automation (IPA) could have three branches: unattended, hybrid, and attended. Decision trees can be used to analyze complex systems and identify the combinations of causal factors that result in a particular outcome.

In the context of configurational causality, decision trees can be used to identify different configurations of causal factors resulting in an outcome. For example, in Chapter 6 and Chapter 7, the four measures of outcome (full-time equivalent (FTE) reduction, process efficiency, process accuracy, and overall IPA success) for successful IPA implementation are discussed. The causal factors that contribute to the outcomes are as follows: bottom-up automation approach, unattended automation type, citizen automation execution, and standalone architecture.

The specific combinations of factors that are most strongly associated with the outcome of an IPA implementation can be identified by exploring different configurations of these causal factors in the decision tree.

This chapter focuses on establishing the causality and how each predictor is associated with the potential outcome of successful IPA implementation.

9.2 Purpose

This study aims to validate the outcomes obtained through configurational causality using decision trees with the potential outcomes view of causality (Mithas et al., 2022). The potential outcomes view of causality is a framework for understanding the causality commonly used in statistics and social sciences. At the heart of this framework is the concept of a "potential outcome," which refers to the outcome that would be achieved if a particular treatment or intervention was administered to an individual; in our case, it is IPA implementations in banking and financial services.

The average treatment effect (ATE) is a measure of the overall impact of a treatment on a population. It is calculated by calculating the difference between the average outcome for the treated group and the average outcome for the control group, which can be mathematically expressed as follows:

$$ATE = E[Y(1)] - E[Y(0)]$$

where Y (1) is the potential outcome under treatment, Y (0) is the potential outcome under control, and E [] represents the expected value. In this study, the before and after analysis is used where Y (1) is the potential outcome before IPA automation and Y (0) is the potential outcome after automation. The ATE is useful because it provides an approach to quantify the overall impact of a treatment on a population.

Overall, the potential outcomes view of causality using the ATE provides a framework for understanding the causal impact of treatments or interventions on populations. By comparing the outcomes between treated and untreated individuals, researchers can determine the overall impact of a treatment and make informed decisions about its implementation. In this study, the ordinary least square (OLS) method of linear regression is used to quantify the effect on the outcomes of IPA based on the configurations derived from decision tree induction.

9.3 Data and Variables

The input to the econometric model is the data of 176 live IPA implementations used for decision tree induction and the results obtained based on the top predictors indicating the success of IPA implementation. These data are prepared in two time periods, i.e., before IPA implementation and after IPA implementation (before and after study), determined by distinct IPA implementation. A total of 352 business process time period observations are considered, of which 176 are before automation and 176 are after automation.

Two dependent variables, namely FTE reduction and process efficiency, and three independent variables—automation type, automation approach, and automation execution— are examined in this study based on decision tree analysis discussed in Chapter 6 and Chapter 7. The remaining eight variables are treated as control variables. Chapter 5 (measures) provides detailed explanations of all thirteen variables. Table 30 presents the definitions and assigned values for performing the before–after analysis using ordinary least squares (OLS).

Table 31: Description and Measurement of Variables

Variable	Description
Dependent variables	
FTE (full-time equivalent) reduction	FTE reduction, or full-time equivalent reduction, is a measure of the number of full-time employees who can be replaced by automation or other efficiency measures.
Process efficiency	Process efficiency is being able to take less time to do things or being able to do more within the same amount of time.
Independent variables	
Automation type	Automation type depends on the extent of the manual intervention when a process is automated. The base variable in this study is attended=0, hybrid=1, and unattended=2
Automation approach	Automation approach is defined in two ways: top-down (base variable = 0) and bottom-up (base variable=1)
Automation execution	Those executing the automation to run the process form the part of automation execution. In general, the execution of the process can be triggered in two ways: citizen automation (base variable=0) and technology-driven automation (base variable=1)
Control variables	
Build vs buy	Decision to build or buy the solution to execute IPA: base variable=0, both build and buy=1, and only buy=2
Domain category	Domain category is a categorical variable, simply a string determining which domain the process belongs to, i.e., asset management, F&A, etc.
Key processes	These are specific processes under the domains represented as a categorical variable
Process complexity	The complexity of the business process being automated: base variable low=0, medium=1, and high=2
Architecture	This represents how the IPA architecture is defined and how it affects the implementation: base variable distributed architecture=0 and standalone architecture=1
Artificial intelligence	This variable is to see the effect of AI on IPA implementation: base variable no=0 and yes=1
Interoperability	This refers to the ability of different RPA systems to work together seamlessly, allowing them to share data and processes across different platforms: base variable no=0 and yes=1
Coding feature	This explains how complex was the coding process: base variable low=0, medium=1, and high=2

9.4 Econometric Specification

To analyze these data, a before–after analysis is conducted using the OLS method (Hayes and Matthes, 2009). OLS is used in linear regression analysis to evaluate the parameters of a linear equation that establishes the relationship between a dependent variable and one or more independent variables. In this study, FTE reduction and process efficiency are considered dependent variables, whereas automation type, automation approach, and automation execution are considered independent variables. OLS minimizes the sum of the squared differences between the observed values of the dependent variable and the values predicted by the linear equation. It generates estimates for the coefficients (slopes) and the intercept of the linear equation that provides the best fit to the data.

To apply OLS, one must first specify a linear model that describes the relationship between the dependent variable and the independent variables, which can be written as follows:

$$Y = \beta 0 + \beta 1X1 + \beta 2X2 + \cdots \beta kXk + \epsilon +$$

where Y is the dependent variable, X1, X2, … Xk are the independent variables, β0 is the intercept, β1, β2, … βk are the coefficients or slopes, and ε is the error term.

There could be heteroscedasticity in this model. It is a statistical term used to describe the scenario where the variability of a dependent variable is unequal across the range of values of an independent variable. In other words, the variance of errors or residuals in a regression model is not constant for all values of the predictor variable(s).

In regression analysis, heteroscedasticity can lead to biased and inefficient estimates of regression coefficients and can result in incorrect inferences about the relationship between the independent and dependent variables. There are several methods to detect and correct heteroscedasticity, such as transforming the dependent or independent variables, using weighted least squares, and applying robust standard errors or white robust/sandwich estimators of standard errors. It is important to address the heteroscedasticity before drawing conclusions from a regression model.

In this study, a white robust/sandwich estimator of standard errors (also known as white-corrected or heteroscedasticity-consistent) is used to correct for heteroscedasticity in the data in regression analysis. The traditional standard errors calculated for regression models assume that the variance of the errors is constant across all observations. However, if there exists heteroscedasticity, these standard errors are biased and may lead to incorrect conclusions about the significance of the estimated coefficients.

White robust standard errors adjust for heteroscedasticity by estimating the variance–covariance matrix of the errors using a modified version of the residual sum of squares. This method considers different variances of errors across the range of values of the independent variables.

White robust standard errors are useful when the assumption of constant variance in errors is violated, and they can provide more reliable estimates of standard errors, t-statistics, and p-values in regression models. They are commonly applied in econometrics and other fields where heteroscedasticity of data is a common feature.

To investigate the predictors or critical success factors of IPA, this research involves a model with two dependent variables (FTE reduction and process efficiency) and three additional dependent variables derived from the analysis of four outcome predictors of IPA identified at significant tree levels. This can be expressed as follows:

$$DV = f(Automation\ Type, Automation\ Approach, Automation\ Execution, Automation\ Type * Automation\ Approach, Automation\ Type * Automation\ Execution, Automation\ Approach * Automation\ Execution)$$

A before–after analysis for the dependent variables FTE reduction and process efficiency is carried out to understand the impact of the independent variables derived from decision tree induction in previous chapters. The before–after analysis is a type of evaluation method used to determine the effect of an intervention or treatment on a particular outcome or a combination of outcomes (in this study, the treatment being IPA implementation). It involves comparing the state of the outcome or outcomes of interest before the treatment is implemented (the "before" period) to the state of the outcome or outcomes after the treatment is implemented (the "after" period).

The before–after analysis aims to assess whether the intervention or treatment has a significant effect on the outcome or outcomes of interest. It is often used in program evaluation, healthcare research, and other fields where determining the effectiveness of an intervention or treatment is important.

In this study, a hierarchical approach of regression is used. In regression, it refers to a method of building regression models by adding predictors in a stepwise manner based on their importance in explaining the variation in the response variable. This approach is often used when many predictors are available for inclusion in the model. It involves fitting a series of models, each with a diverse set of predictors. The first model includes only the most important predictor determined based on expert knowledge or prior research. In the subsequent models, additional predictors are added in a stepwise manner, with each predictor being evaluated for its ability to improve the overall fit of the model.

The hierarchical approach can be useful in scenarios where the knowledge about which predictors are most important is limited or when the number of potential predictors is large. By building models in a stepwise manner, this approach can help identify the most important predictors and avoid model overfitting.

In this study, five models are investigated using a hierarchical approach involving three predictors. The first model presents the direct effects on various outcomes; then, in models two to four, the two-way interaction effects of each pair of predictors are studied, respectively, and in the fifth model, the two-way interaction effects of all pairs of predictors are examined together.

9.5 Results

In Table 32, the results of the direct effects of the predictors on the automation approach, automation type, and automation execution on the IPA implementation success outcome variable FTE reduction are presented. The first column of the table represents model 1, which presents the direct effects of the predictors on FTE reduction before and after the automation.

In model 1, before automation, the processes that undergo the bottom-up approach on average show a 6.4% FTE reduction compared with the processes that undergo the top-down approach. After automation, on average, a 22% FTE reduction is observed for the processes that undergo the bottom-up approach compared with the processes that undergo the top-down approach.

In model 1, before automation, the processes that undergo unattended treatment on average show a 20% FTE reduction compared with those undergoing attended treatment. After automation, on average, a 53.3% FTE reduction is observed for the processes with unattended automation compared with those with attended automation.

In model 1, before automation, the processes with citizen execution show no significant difference from those with tech-driven execution. However, after automation, on average, an 11.3% FTE reduction is observed for the processes with citizen automation execution compared with those with tech-driven execution.

The second column of the table represents model 2, which presents the two-way interaction effects of the first pair of predictors, i.e., automation type × automation approach, on the IPA implementation success outcome FTE reduction before and after automation.

In model 2, before automation, the processes with a two-way interaction of the pair with a combination of the bottom-up approach × hybrid treatment and the bottom-up approach × unattended treatment show no significant difference from those with the bottom-up approach × attended treatment. However, after automation, on average, a 6% of FTE reduction is observed for the processes with the bottom-up approach × unattended automation compared with the processes with the bottom-up approach × attended automation.

The third column of the table represents model 3, which presents the two-way interaction effects of the second pair of predictors, i.e., automation approach × automation execution on the IPA implementation success outcome FTE reduction before and after automation.

In model 3, before automation, the processes with a two-way interaction of the bottom-up approach × citizen execution show no significant difference from those with the bottom-up approach × tech-driven execution. However, after automation, on average, a 35% FTE reduction is observed for the processes with the bottom-up approach × citizen execution from those with the bottom-up approach × tech-driven execution.

The fourth column of the table represents model 4, which presents the two-way interaction effects of the third pair of predictors, i.e., automation type × automation execution on the IPA implementation success outcome FTE reduction before and after automation.

In model 4, before and after automation, the processes with the two-way interaction of hybrid treatment × citizen execution and unattended treatment × citizen execution show no significant difference from those with attended treatment × tech-driven execution.

The fifth column of the table represents model 5, which presents the two-way interaction effects of all three pairs of predictors, i.e., automation type × automation approach, automation approach × automation execution, and automation execution × automation type on the IPA implementation success outcome FTE reduction before and after automation.

In model 5, before and after automation, the processes with the two-way interaction of all three pairs show no significant difference.

Table 32: Main Results of Econometrics—FTE Reduction

Dependent Variable FTE Reduction	(1) Direct Effects	(2) Two-Way Interaction Effects	(3) Two-Way Interaction Effects	(4) Two-Way Interaction Effects	(5) Two-Way Interaction Effects
Pre					
Approach Bottom-Up	**-.64 (.22, 0.005)** **	-.35(.66,0.592)	**-.63(.22,0.006)** **	**-.66(.23, 0.005)** **	-.29(1.0,0.784)
Type: Hybrid Unattended	-.496 (0.33,0.134) **-1.99(0.34, 0.000)** ***	,	-.47(.33,0.157) **-1.97(.35,0.000)** ***	-.37(.40, 0.344) **-1.89(.40,0.000)** ***	-.07(.47, 0.876) **-2.0(.46,0.000)** ***
Execution Citizen Bots	-.24(.35, 0.49)	-.16(.35, 0.649)	.42(1.8,0.816)	.21(.75,0.777)	.59(1.81, 0.746)
Approach X Type Bottom-Up X Hybrid Bottom-Up X Unattended	---	-.69(.69,0.315) .05(.71,0.940)	---	---	-.76(1.0,0.483) .01(1.0,0.992)
Approach X Execution Bottom-Up X Citizen Bots	---	---	-0.7(1.8,0.700)	---	-.75(2.13, 0.724)
Type X Execution Hybrid X Citizen Bots Unattended X Citizen Bots	---	---	---	-.48(.65,0.463) -.50(1.12,0.655)	.04(1.06, 1.06) -.29(1.39, 0.835)
Post					
Approach X Post Bottom-Up X Post	**-1.57 (.34, 0.000)** ***	**-2.02(.89,0.025)** **	**-1.5(.34, 0.000)** ***	**-1.58(.35,0.000)** ***	-1.37(1.68,0.415)
Type X Post Hybrid X Post Unattended X Post	-.61(.49,0.213) **-3.76(.50, 0.000)** ***	-.22(.67, 0.740) **-4.47(.66,0.000)** ***	-.56(.49, 0.254) **-3.73(.50,0.000)** ***	-.54(.59,0.367) **-3.73(.59,0.000)** ***	-.06(.68,0.927) **-4.32(.68, 0.000)** ***
Automation Execution X Post Citizen Bots X Post	**-0.89(.45, 0.050)** **	-.61(.45, 0.175)	**1.77(.43,0.000)** ***	-.79(.90, 0.380)	**1.77(.51, 0.001)** ***

Approach X Type X Post Bottom-Up X Hybrid X Post Bottom-Up X Unattended X Post	---	-.50(.34, 0.137) **-.64(.33, 0.054) ***	---	---	-1.00(1.76,0.571) .76(1.75, 0.665)
Approach X Execution X Post Bottom-Up X Citizen Bots X Post	---	---	**-2.79(.58,0.000) ****	---	-2.95(1.84, 0.110)
Type X Execution X Post Hybrid X Citizen Bots X Post Unattended X Citizen Bots X Post	---	---	---	-.28(1.0,0.780) .35(.96, 0.713)	.44(1.85,0.809) .39(1.81, 0.829)
n value - Observations	352	352	352	352	352
Groups	176	176	176	176	176
Time Period	2	2	2	2	2
R Squared	0.86	0.8652	0.861	0.8604	0.8661

Notes
Significance level: P<0.01 ***, P<0.05 ** & P<0.1 *
Base variable for the automation approach is taken as top-down.
Base variable for the automation type is taken as unattended.
Base variable for automation execution is taken as tech-driven automation.
Bottom-up automation approach shows significant reduction in FTE.
"Unattended" automation type shows consistent reduction in FTE.
When automation type is combined with automation approach, a significantly larger effect is observed compared with individual effects, i.e., automation type and automation approach moderate the effect on FTE reduction.
When automation approach is combined with automation execution, a significantly larger effect is observed compared with individual effects, i.e., automation approach and automation execution moderate the effect on FTE reduction.

Table 33 presents the results of the direct effects of the predictors on the automation approach, automation type, and automation execution on the IPA implementation success outcome variable process efficiency. The first column of the table represents model 1, which presents the direct effects of the predictors on process efficiency before and after automation.

In model 1, before automation, the processes with the bottom-up approach show no significant difference from those with the top-down approach. However, after automation, on average, an 8.25% improvement in process efficiency is observed for the processes with the bottom-up approach compared with those with the top-down approach.

In model 1, before automation, the processes with the hybrid treatment on average show an 8.6% improvement in process efficiency, and those with the unattended treatment on average show a 19% improvement in process efficiency compared with those with attended treatment. After automation, on average, a 23.4% improvement in process efficiency is observed for the processes with the hybrid treatment, and on average, a 56.5% improvement in process efficiency is observed in those with the unattended automation compared with those with the attended automation.

In model 1, before automation, the processes with citizen execution show no significant difference from those with tech-driven execution. However, after automation, on average, a 9.5% reduction in process efficiency is observed for the processes with citizen automation execution compared with the processes with tech-driven execution.

The second column of the table represents model 2, which presents the two-way interaction effects of the first pair of predictors, i.e., automation type × automation approach on the IPA implementation success outcome process efficiency before and after automation.

In model 2, before and after automation, the processes with a two-way interaction of the pair with the combination of bottom-up approach × hybrid treatment and bottom-up approach × unattended treatment show no significant difference from those with the bottom-up × attended treatment.

The third column of the table represents model 3, which presents the two-way interaction effects of the second pair of predictors, i.e., automation approach × automation execution on the IPA implementation success outcome process efficiency before and after automation.

In model 3, before automation, the processes with a two-way interaction of the pair with the combination of bottom-up approach × citizen execution show no significant difference from those with bottom-up approach × tech-driven execution. However, after automation, on average, a 56.6% reduction in process efficiency is observed for the processes with bottom-up approach × citizen execution.

The fourth column of the table represents model 4, which presents the two-way interaction effects of the third pair of predictors, i.e., automation type × automation execution on the IPA implementation success outcome process efficiency before and after automation.

In model 4, before and after automation, the processes with a two-way interaction of the pair with the combination of hybrid treatment × citizen execution and unattended treatment × citizen execution show no significant difference from those with attended treatment × tech-driven execution.

The fifth column of the table represents model 5, which presents the two-way interaction effects of all three pairs of predictors, i.e., automation type × automation approach, automation approach × automation execution, and automation execution × automation type on the IPA implementation success outcome process efficiency before and after automation.

In model 5, before automation, the processes with the two-way interactions of all three pairs of predictors show no significant difference compared with the respective base variables. However, after automation, on average an 18.5% improvement in process efficiency is observed with bottom-up approach × hybrid treatment and a 20% improvement with bottom-up approach × unattended treatment compared with top-down approach × attended treatment. When the processes receive two interactions of the pair of predictors bottom-up approach × citizen execution, there is, on average after automation, a 90% reduction in process efficiency compared with top-down approach × tech-driven execution. Finally, after automation, on average, a 37.4% improvement in process efficiency is observed with hybrid treatment × citizen execution, and a 47.2% improvement in process efficiency is observed with unattended treatment × citizen execution compared with attended treatment × tech-driven execution.

Table 33: Main Results of Econometrics Process Efficiency

Dependent Variable FTE Reduction	(1) Direct Effects	(2) Two-Way Interaction Effects	(3) Two-Way Interaction Effects	(4) Two-Way Interaction Effects	(5) Two-Way Interaction Effects
Pre					
Approach Bottom Up	-1.22(1.05,0.247)	-1.24(4.34, 0.775)	-13.72(11.80,0.246)	-1.18(1.07,0.269)	-17.29(12.13,0.155)
Type Hybrid Unattended	**-5.16(1.8,0.005) ***** **-11.36(1.96, 0.000) *****	**-5.63(2.34,0.017) ***** **-11.03(2.40, 0.000) *****	**-4.74(1.83,0.010_**** **-10.94(1.96, 0.000) *****	-4.21(3.37, 0.213) **-9.85(5.31,0.065) ****	4.63(7.84, 0.555) 1.27(8.58,0.882)
Execution Citizen Bots	.58(1.9, 0.763)	.71(1.93,0.714)	-11.23(11.80,0.342)	1.89(4.97,0.704)	-10.74(11.90, 0.367)
Approach X Type Bottom-Up X Hybrid Bottom-Up X Unattended	---	.64(4.25,0.880) -.58(4.44,0.895)	---	---	-5.79(7.46, 0.438) -6.95(7.43,0.350)
Approach X Execution Bottom-Up X Citizen Bots	---	---	12.64(11.80,0.285)	---	22.52(14.03, 0.109)
Type X Execution Hybrid X Citizen Bots Unattended X Citizen Bots	---	---	---	-1.24(4.08,0.760) -1.80(5.93, 0.761)	-5.79(7.46,0.438) -6.95(7.43,0.350)
Post					
Approach X Post Bottom-Up X Post	**-3.73(1.71, 0.030) *****	-7.27(5.65, 0.199)	**-24.59(3.25, 0.000) *****	**-3.46(1.75, 0.050) ****	**-27.25(4.54, 0.000) *****
Type X Post Hybrid X Post Unattended X Post	**-8.89(2.7,0.001) ***** **-22.54(2.73, 0.000) *****	**-10.78(3.57, 0.003) ***** **-23.90(3.49, 0.000) *****	**-8.51(2.71, 0.002) ***** **-22.27(2.74, 0.000) *****	-5.77(5.13, 0.261) **-15.07(5.39,0.005) *****	7.17(7.66,0.350) -1.25(7.78,0.872)
Execution X Post Citizen Bots X Post	**5.09(2.69, 0.060) ***	4.12(2.89, 0.155)	**-15.31(2.44, 0.000) *****	8.66(5.49, 0.116)	**-14.74(2.88, 0.000) *****
Approach X Type X Post Bottom-Up X Hybrid X Post Bottom-Up X Unattended X Post	---	4.33(5.92,0.465) 3.64(6.00,0.545)	---	---	**-11.12(5.72, 0.053) *** **-11.99(5.52,0.031) ****
Approach X Execution X Post Bottom-Up X Citizen Bots X Post	---	---	**21.33(3.36, 0.000) *****	---	**35.39(6.80, 0.000) *****
Type X Execution X Post Hybrid X Citizen Bots X Post	---	---	---	-4.92(6.04, 0.416) -9.26(6.23,0.138)	**-16.70(6.77, 0.014) **** **-21.39(6.96, 0.002) *****

Unattended X Citizen Bots X Post					
n value - Observations	352	352	352	352	352
Groups	176	176	176	176	176
Time Period	2	2	2	2	2
R Squared	0.9116	0.9119	0.9138	0.912	0.9164
Notes Base variable for the automation approach is taken as top-down. Base variable for the automation type is taken as unattended. Base variable for automation execution is taken as tech-driven automation. Bottom-up automation approach shows significant improvement in process efficiency. "Unattended" automation type shows consistent improvement in process efficiency. "Citizen bots" automation execution decreases process efficiency. When automation type is combined with automation approach, a significantly larger positive effect is observed compared with individual effects, i.e., automation type and automation approach moderate the effect on process efficiency. When automation approach is combined with automation execution, a significantly negative effect is observed compared with individual effects, i.e., automation approach and automation execution moderate the effect on process efficiency negatively.					

9.6 Implications

In Chapter 6, the decision tree induction on the live data of 176 IPA implementations and from the tree analysis is discussed, and sixteen rules are derived. In this chapter, the OLS method of linear regression is used to quantify the effect on the outcomes of IPA based on the configurations derived from decision tree induction and to validate the rules through econometric analysis.

In Table 32, the results of the econometric analysis of predictors affecting IPA outcome FTE reduction before and after automation are presented, which validate the rules derived in Chapter 6 for FTE reduction. From these results, on average, the results of unattended automation are significant, with a 53.3% FTE reduction after automation compared with attended automation. This validates rule 1, which predicts high FTE reduction, and rule 3, which predicts low FTE reduction. Similarly, on average, the results of the bottom-up approach are significant, with a 22% FTE reduction after automation compared with the top-down approach. This validates rule 2.

In Table 33, the results of the econometric analysis of predictors affecting the IPA outcome process efficiency, before and after automation, are presented, which validate the rules derived in Chapter 6 for process efficiency. From these results, on average, hybrid automation results in a 23.4% improvement in process efficiency after automation compared with attended, which validates rule 4 and rule 7. Next, on average, unattended automation results are significant, with a 56.5% improvement in process efficiency compared with attended automation, which validates rule 5, rule 6, and rule 8. With respect to the interaction effects on process efficiency, it is clear that after automation, on average, a 37.4% improvement in process efficiency is observed with hybrid treatment × citizen execution, which validates rule 4.

Overall, from the results of the econometric analysis, the potential outcome causality is established with the configurational view of causality established by decision trees.

9.7 Conclusion

Thus far, in this chapter, the application of econometric analysis in validating the sixteen rules and six propositions derived in previous chapters is discussed. The specifications for our analysis are defined in terms of dependent, independent, and control variables, and the before–after analysis is carried out using OLS.

The results establish the validity of the rules and propositions obtained using the decision tree induction; hence, the configurational causality is validated by potential outcome-based causality. The configurational causality using decision trees determines what are the critical success factors or predictors affecting the success of IPA implementation. The potential outcome causality using OLS not only establishes the outcomes of configurational causality but also quantifies the effects of the factors affecting the success of IPA implementation. In the next chapter, the theoretical contributions, managerial implications, strengths, and limitations are discussed along with the concluding thoughts.

10 DISCUSSION

This chapter examines the sixteen rules, six insights, and six meta-propositions that affect the successful implementation of intelligent process automation (IPA) and discusses the theoretical and managerial implications, strengths, and limitations of this study, besides shedding light upon future research opportunities.

10.1 Theoretical Contributions

The results of this study are consistent with the ongoing discussion on identifying critical factors for the successful implementation of IPA. In this study, predictors that have a significant effect on the success of IPA implementation are identified. For instance, unattended automation and bottom-up approach are important predictors that appear in all four outcomes of successful IPA implementation, whereas attended automation is a significant predictor in all outcomes of unsuccessful IPA implementation. These findings suggest that dominant predictors that significantly affect the successful implementation of IPA are purely strategic profiles (Kathuria et al., 2020).

In addition, this study explains how software projects are different from IPA implementations. In this research, predictors of IPA implementation success that are distant from software project predictors are identified, thus significantly contributing to the theory.

Furthermore, this study contributes to moving the conversation forward on how to ensure the successful implementation of technology democratization based on artificial intelligence (AI).

Besides the implications of decision rules, insights, and propositions, this research provides three significant contributions to theory. First, decision trees provide a comprehensive and easily comprehensible representation of theories. These trees categorize and sort predictors according to their importance on multiple levels and remove non-informative predictors, thus allowing researchers to focus on informative ones. In addition, decision trees organize predictors based on their importance, providing a summary of the decision-making process and experiences of the decision-makers. Second, rules extracted from the decision trees reveal ontologies or concepts and categories that define their properties and relationships. Hence, this study demonstrates how decision trees can illuminate first principles or "the essence of things," which is a major contribution to theory.

Third, this study highlights the use of abduction as a logical framework for theory development. Two commonly used approaches in logical conclusions are deduction and induction, with the former starting from a known rule and seeking to apply it to a case to obtain knowledge (Reichertz, 2007) and the latter beginning with a case and extending a result from the data into a rule. In contrast, in abduction (Hobbs et al., 1993), the best explanation is inferred from the available information. In this study, theoretically nuanced explanations were developed, and their predictions were refined by extracting rules from decision trees. Through abstraction, the rules were examined in their abstract forms, thus enabling them to progressively develop and refine their theoretical propositions. In this approach, the focus was not on actual rule instantiations, but on discovering how abstractions progressively evolved into an ontology of rules. This approach provided a better understanding of the combined influence of predictors on the outcome of interest.

In other words, abductive discovery is used to proceed from the data (four trees), to rules (sixteen context-specific rules and their general forms), to mid-level theoretical insights(six insights), and to finally arrive at cases (six generic propositions) (Reichertz, 2007). Herein lies our third significant contribution.

Next, an alternative means for developing multilevel theory that reveals sequences of insights is demonstrated. Hierarchical linear modeling is used in extant efforts at multilevel theory building (e.g., (Maruping and Magni, 2012, Suh et al., 2011), and multistage econometric models (e.g., (Xie and Lee, 2015) are used to examine the phenomena between and across two levels of analysis. This distinct methodological contribution lies in our articulation of sequences of predictor combinations that lie across the four levels of theory, i.e., governance, process, technology, and complexity levels. The order in the decision tree induction goes beyond mere outcomes by explicitly representing these sequences, where the partial ordering of decisions is of importance.

Third, a combination of predictors that would enable organizations to achieve specific outcomes of IPA implementation is identified, namely full-time equivalent (FTE) reduction, process efficiency, and accuracy.

Overall, this manuscript moves Information Systems (IS) research forward by presenting an alternative form of knowledge production that emphasizes "inductive, rich inquiries using innovative and extensive data sets" and enables "novel, genuine, high-level theorizing around germane conceptual relationships" (Grover and Lyytinen, 2015). Our data-driven, abductive–inductive–abductive research interprets the patterns in data to discover empirical regularities (stylized facts) that challenge the existing beliefs and give rise to new constructs and theories (Helfat, 2007). As Weick (Weick, 1995) notes, novel theories require diverse lenses to examine the phenomena present in them, keen observation of data, disciplined imagination, and thought experiments. This manuscript presents an alternate lens that can help IS theorists develop rich theory by seamlessly moving through different levels of abstractions to discover new knowledge and ontologies and to identify inter-relationships across facts (Grover and Lyytinen, 2015).

10.2 Managerial Implications

This study investigates the success of IPA implementation from the vantage point of various predictors (e.g., automation type, automation approach, automation execution, etc.). Thus, the implications of the findings of this study are multifold based on sixteen rules, six insights, and five propositions that are identified following managerial implications.

- A nuanced view into the decision-making process for IPA practitioners is provided regarding the predictors or critical factors affecting both high and low success of IPA implementation, specifically helping IPA practitioners to determine factors that contribute to high automation success and those that do not. These insights help derive rational decision-making mechanisms by IPA practitioners/managers regarding critical factors that determine the successful implementation of IPA.
- For successful implementation of IPA, organizing principles for their implementations are identified by highlighting the most efficient path for successful IPA implementation, thus improving the probability of success. Following Rules 1–16, IPA managers should focus on unattended automation, i.e., they need to choose the right business process to undergo automation so that IPA implementation is understood well in advance and made seamless without human intervention. Then, the bottom-up automation approach needs to be focused on, especially in the combination of hybrid or unattended automation type. By focusing on these two critical factors, the probability of successful IPA implementation becomes high.
- Before implementing IPA, it is essential to define what the organization wants to achieve through automation (outcomes of measure). This process includes the following: setting specific goals and objectives, identifying key performance indicators (KPIs) to measure the success, and aligning IPA initiatives with the overall strategy of the organization. In this study, the outcomes FTE reduction, process efficiency, accuracy, and IPA success are discussed, depending upon the organizational priorities.

- Not all processes are suitable for automation. Therefore, repetitive, rule-based, and high-volume processes need to be identified and evaluated for automation. This will ensure that the organization maximizes the benefits of IPA while minimizing the risks of implementation.
- This study presents a well-defined IPA strategy that can help organizations achieve their automation goals. This strategy includes a roadmap for implementation, a clear timeline for deployment, and an assessment of the effects on the workforce.

Practice implications also extend to other contexts of automation implementations such as low/code in healthcare, retail, and other sectors.

10.3 Strengths and Limitations

This research has several strengths. Live sample data of 176 real-time implementations are used to systematically examine critical factors that lead to successful IPA implementation. This sample is specifically focused on banking and financial services across the world, thus taking into account complex issues and the heterogeneity of the data.

This research setup addresses the potential concerns in sample selection as both successful and not-so-successful implementations are part of the dataset. This comprehensive dataset contributes unique insights into successful IPA implementation, especially for complex processes in banking and financial services across the world. Second, the decision tree induction methodology used in this study also contributes unique insights by identifying key patterns in the data and presents the analytics in an intuitive and easy-to-follow manner for a wide variety of stakeholders. Critically, this methodology is appropriate for inferring a fit (or a lack of fit) between what managers are expected to do (in theory) and what they actually do (as revealed by the trees), making it appropriate for presenting to managers (Drazin and Van de Ven, 1985). Third, the decision tree induction methodology has a low rate of false-positive predictions (Spangler et al., 1999). Thus, our low prediction error of ~33% is also conservative in nature. Finally, a key strength of this study lies in the use of decision tree induction to realize a new ontology, which signals future research advances. This methodology offers a perspective that aligns with the current trends in AI such as semantic networks and cognitive computing (Davenport and Ronanki, 2018), which require explicit representation of extensive knowledge. Decision trees create ontologies that can be used to further derive semantics and knowledge representations. Therefore, ontologies are the pillars of the semantic web that enable us to understand first principles, or "the essence of things." If organizations are conceived as bundles of decisions dynamically allocated across humans, systems, or combinations of humans and systems, decision trees represent the first and vital step toward achieving cognitive reapportionment and autonomous decision-making (Konsynski and Sviokla, 1993). Eventually, decision trees could credibly approximate the decision processes and governing dynamics pertinent to management practices.

There are some limitations to the present study. First, decision trees are an approximation, albeit credible, of the decision-making process in identifying the critical factors for successful IPA implementation. Though follow-up interviews with IPA practitioners are conducted to ascertain the validity of the results, this study cannot precisely quantify the exact order of steps taken by managers in the decision-making process for identifying the critical factors for the successful implementation of IPA. Second, only a sample of IPA implementations in banking and financial services is discussed. A larger sample would help us to ascertain the propositions made in this study. Third, the generalizability of these results to other types of technology automation may be limited. Finally, the cross-sectional nature of our data precludes us from drawing causal conclusions through our analysis. Because of the presence of different approaches to IPA implementation in the market across domains and the ongoing technological advances especially due to AI, gathering and studying longitudinal effects in this context is not feasible. However, this is an interesting scope for future research studies in other contexts.

10.4 Directions for Future Research

Two avenues for future research are identified in this study. First, our research methodology utilizes induction to discover rules explaining the critical factors for the successful implementation of IPA. Rules serve as the primary input to the abduction process. Thus, this methodology serves as a harbinger for cognitive computing, whereby AI systems can mimic the functioning of the human brain and help improve human decision-making by inferring the best explanation from a given set of rules. A further step toward this goal can be made by mapping decision journeys. Decision-tree-based abduction, which helps define ontologies and construct decision journeys and flows, is thus a stepping stone toward a deeper understanding of decision-making by human agents (stakeholders) in complex situations. Future research for automating this intellectual improvement from rules to cases, or inference to the best explanation, is a foundational feature to realize the dream of cognitive computing.

Second, in this study, IPA implementation is studied from the point of view of a few success predictors. Future research could contrast related questions from the perspectives of other technological advancements such as AI, augmented reality/virtual reality, and right process selection variables for successful IPA implementation.

10.5 Summary of Key Findings

This section presents the summary of key findings of sixteen rules, six insights, and six propositions that immensely contribute to the theory for IS research, as shown in Table 33.

Table 34: Summary of Theoretical Artifacts

Theoretical Artifact
Rules
Rule 1: Unsupervised intelligence—if the automation type is unattended, then FTE reduction is high.
Rule 2: Empowering business user—if the automation type is hybrid and the automation approach is bottom-up, the FTE reduction is high.
Rule 3: Partial intelligence "Guide Me"—if the automation type is attended, then FTE reduction is low.
Rule 4: Citizen intelligence and standalone systems—if the automation type is hybrid and executed by citizen, then process efficiency is high for standalone systems.
Rule 5: Partial intelligence "Guide Me"—if the automation type is attended, then process efficiency is low.
Rule 6: Enterprise automation—if the automation type is unattended and when systems are interoperable, process efficiency is high with the top-down approach.
Rule 7: Citizen intelligence and distributed systems—if the automation type is hybrid and executed by citizen, then process efficiency is low for distributed systems.
Rule 8: Unsupervised intelligence—if the automation type is unattended, then process efficiency is high.
Rule 9: Empowering business user—if the automation approach is bottom-up and when the automation type is either hybrid or unattended, then process accuracy is high.
Rule 10: Disenfranchise business user automation—if the automation approach is top-down and when the automation type is either hybrid or attended, then process accuracy is low.
Rule 11: Enterprise automation—if the automation approach is top-down and the automation type is unattended, then process accuracy is high.
Rule 12: Unsupervised intelligence—if the automation type is unattended, then overall IPA success is high.
Rule 13: Average intelligence "Tackling Complexity"—if the automation type is hybrid and complexity is medium, then overall IPA success is moderate.
Rule 14: Partial intelligence "Guide Me"—if the automation type is attended, then overall IPA success is low.
Rule 15: Empowering business owner—if the automation type is hybrid and complexity of the business process is high, then overall IPA success with the bottom-up automation approach is high.
Rule 16: Disenfranchise business owner: If business processes are highly complex and have a hybrid automation type, then overall IPA success is low with the top-down automation approach.
Insights
Insight 1: Unattended automation type is a necessary predictor for the top-ranked rule that results in high IPA implementation success for all outcomes.
Insight 2: Attended automation type is a necessary predictor for the top-ranked rule that results in low IPA implementation success for all outcomes.
Insight 3: Bottom-up automation approach is a necessary predictor for the second-ranked rule that results in high IPA implementation success for the majority of outcomes; however, it does not contribute to IPA implementation outcome process efficiency.
Insight 4: Top-down automation approach is an important predictor that results in high IPA implementation success along with specific combinations of other predictors for the majority of outcomes; however, it does not contribute to IPA implementation outcome FTE reduction.

Insight 5: Citizen automation execution predictor results in high process efficiency when combined with standalone architecture and hybrid automation; however, it does not contribute to any of the other IPA implementation success outcomes.
Insight 6: Citizen automation execution predictor results in low process efficiency when combined with distributed architecture and hybrid automation; however, it does not contribute to any of the other IPA implementation success outcomes.
Propositions
Proposition 1: Predictors determining the success of IPA implementation are present across multiple levels of analysis.
Proposition 2: Combinations of predictors determining the high success of IPA implementation are present across multiple levels of analysis and are different from the predictors informing low success of IPA implementation.
Proposition 3: Predictors determining high or low success of IPA implementations at one level influence the inclusion or exclusion of predictors at the same or other levels.
Proposition 4: Multiple predictors determining the success of IPA implementation may interact such that a few predictors either dominate or reinforce their combined influence on IPA success.
Proposition 5: One or a combination of predictors determining overall IPA success may or may not affect other specific outcomes of IPA success.
Proposition 6: If the presence of a predictor is necessary for a successful IPA implementation outcome, the opposite does not imply that it will lead to failure.

10.6 Concluding Thoughts

This study offers three key takeaways for researchers, managers, and practitioners. First, it provides rules, insights, and propositions to identify the dominant predictors and combination predictors present at all theoretical levels explaining successful IPA implementation. Second, it discusses the presence of interdependencies between predictors of IPA implementation success outcomes that lie across multiple levels of theory. This study is an attempt to reconcile multiple multilevel predictors. However, because of the complexity, diversity, and uncertainties associated with diverse types of IPA implementation, there may be other combinations of predictors that offer fresh research opportunities. This study calls for further multilevel research that goes beyond examining hierarchical relationships across only two levels of analysis to develop a deeper understanding of successful IPA implementation. Finally, the methodological contributions of this study provide opportunities for developing a richer agenda for IS researchers. Although decision tree induction is not a new tool, advances in the methodology and the availability of large datasets allow researchers to realize its potential. As envisioned by prior research, this methodology allows us to study variables across multiple levels of theory and discover their emergent and tacit combinations. Bringing this approach to nascent, emerging areas of study, such as success factors of IPA implementation, will advance our community, research, and management practice forward on the arc of progress.

11 REFERENCES

AGOSTINELLI, S., MARRELLA, A. & MECELLA, M. Research challenges for intelligent robotic process automation. International Conference on Business Process Management, 2019. Springer, 12-18.

AGOSTINELLI, S., MARRELLA, A. & MECELLA, M. 2020. Towards intelligent robotic process automation for BPMers. *arXiv preprint arXiv:2001.00804*.

ALTER, S. 1978. Development patterns for decision support systems. *Mis Quarterly*, 33-42.

ASATIANI, A., MALO, P., NAGBØL, P. R., PENTTINEN, E., RINTA-KAHILA, T. & SALOVAARA, A. 2020. Challenges of explaining the behavior of black-box AI systems. *MIS Quarterly Executive,* 19**,** 259-278.

ASATIANI, A. & PENTTINEN, E. 2016. Turning robotic process automation into commercial success–Case OpusCapita. *Journal of Information Technology Teaching Cases,* 6**,** 67-74.

AUTH, G., CZARNECKI, C. & BENSBERG, F. Impact of robotic process automation on enterprise architectures. INFORMATIK 2019: 50 Jahre Gesellschaft für Informatik–Informatik für Gesellschaft (Workshop-Beiträge), 2019. Gesellschaft für Informatik eV.

AXMANN, B. & HARMOKO, H. Robotic process automation: An overview and comparison to other technology in industry 4.0. 2020 10th International Conference on Advanced Computer Information Technologies (ACIT), 2020. IEEE, 559-562.

BAILEY, A. & NGWENYAMA, O. 2014. Ethnographic Decision Tree Modelling: An Exploration of Telecentre Usage in the Human Development Context. *Advances in Research Methods for Information Systems Research: Data Mining, Data Envelopment Analysis, Value Focused Thinking***,** 63-77.

BAIYERE, A., SALMELA, H. & TAPANAINEN, T. 2020. Digital transformation and the new logics of business process management. *European Journal of Information Systems,* 29**,** 238-259.

BENBYA, H., DAVENPORT, T. H. & PACHIDI, S. 2020. Artificial intelligence in organizations: current state and future opportunities. *MIS Quarterly Executive,* 19.

BENBYA, H., PACHIDI, S. & JARVENPAA, S. 2021. Special issue editorial: artificial intelligence in organizations: implications for information systems research. *Journal of the Association for Information Systems,* 22**,** 10.

BERENTE, N., GU, B., RECKER, J. & SANTHANAM, R. 2021a. Managing artificial intelligence. *MIS quarterly,* 45.

BERENTE, N., GU, B., RECKER, J. & SANTHANAM, R. 2021b. Managing artificial intelligence. *MIS quarterly,* 45**,** 1433-1450.

BOONSTRA, A. 2003. Structure and analysis of IS decision-making processes. *European Journal of Information Systems,* 12**,** 195-209.

BRÉZILLON, P., PASQUIER, L. & POMEROL, J.-C. 2002. Reasoning with contextual graphs. *European Journal of Operational Research,* 136**,** 290-298.

BYGSTAD, B. & ØVRELID, E. 2020. Architectural alignment of process innovation and digital infrastructure in a high-tech hospital. *European Journal of Information Systems,* 29**,** 220-237.

CARDEN, L., MALDONADO, T., BRACE, C. & MYERS, M. 2019. Robotics process automation at TECHSERV: An implementation case study. *Journal of Information Technology Teaching Cases,* 9**,** 72-79.

CHAIMANKONG, B., CHAIMANKONG, M., JERMSITTIPARSERT, K. & CHETTHAMRONGCHAI, P. 2021. THE IMPACT OF COVID-19 ON THAI CONSUMER BEHAVIOUR IN USING DIGITAL BANKING. *Academy of Strategic Management Journal***,** 1-9.

CHANIAS, S., MYERS, M. D. & HESS, T. 2019. Digital transformation strategy making in pre-digital organizations: The case of a financial services provider. *The Journal of Strategic Information Systems,* 28**,** 17-33.

CHIN, W. W. 1995. Partial least squares is to LISREL as principal components analysis is to common factor analysis. *Technology studies,* 2**,** 315-319.

CHIN, W. W. 1998. Commentary: Issues and opinion on structural equation modeling. JSTOR.

CHOI, D., R'BIGUI, H. & CHO, C. Robotic process automation implementation challenges. Proceedings of International Conference on Smart Computing and Cyber Security: Strategic Foresight, Security Challenges and Innovation (SMARTCYBER 2020), 2021. Springer, 297-304.

CHUNG, M. S. & LEE, J. Y. 2018. Systemic analysis of research activities and trends related to artificial intelligence (aI) technology based on latent dirichlet allocation (lda) model. *Journal of the Korea Industrial Information Systems Research,* 23**,** 87-95.

COOMBS, C., HISLOP, D., TANEVA, S. K. & BARNARD, S. 2020. The strategic impacts of Intelligent Automation for knowledge and service work: An interdisciplinary review. *The Journal of Strategic Information Systems,* 29**,** 101600.

COOPER, L. A., HOLDERNESS JR, D. K., SORENSEN, T. L. & WOOD, D. A. 2019. Robotic process automation in public accounting. *Accounting Horizons,* 33**,** 15-35.

COUNIHAN, A., FINNEGAN, P. & SAMMON, D. 2002. Towards a framework for evaluating investments in data warehousing. *Information Systems Journal,* 12**,** 321-338.

DATTA, P., WALKER, L. & AMARILLI, F. 2020. Digital transformation: Learning from Italy's public administration. *Journal of Information Technology Teaching Cases,* 10**,** 54-71.

DAVENPORT, T. & RONANKI, R. 2018. A real-world guide to artifical intelligence. *Harvard Business Review,* 96**,** 108-116.

DENAGAMA VITHARANAGE, I. M., BANDARA, W., SYED, R. & TOMAN, D. An empirically supported conceptualisation of robotic process automation (RPA) benefits. Proceedings of the 28th European Conference on Information Systems (ECIS2020), 2020. Association for Information Systems.

DIAMANTOPOULOS, A. & WINKLHOFER, H. M. 2001. Index construction with formative indicators: An alternative to scale development. *Journal of marketing research,* 38**,** 269-277.

DRAZIN, R. & VAN DE VEN, A. H. 1985. Alternative forms of fit in contingency theory. *Administrative science quarterly***,** 514-539.

EDWARDS, J. R. & BAGOZZI, R. P. 2000. On the nature and direction of relationships between constructs and measures. *Psychological methods,* 5**,** 155.

EL SAWY, O. A., MALHOTRA, A., PARK, Y. & PAVLOU, P. A. 2010. Research commentary—seeking the configurations of digital ecodynamics: It takes three to tango. *Information systems research,* 21**,** 835-848.

ENGEL, C., ELSHAN, E. & EBEL, P. 2021. Moving Beyond Rule-Based Automation: A Method for Assessing Cognitive Automation Use Cases.

FÜGENER, A., GRAHL, J., GUPTA, A. & KETTER, W. 2021. Cognitive challenges in human–artificial intelligence collaboration: investigating the path toward productive delegation. *Information Systems Research*.

GAMI, M., JETLY, P., MEHTA, N. & PATIL, S. Robotic Process Automation–Future of Business Organizations: A Review. 2nd International Conference on Advances in Science & Technology (ICAST), 2019.

GORWA, R. & GUILBEAULT, D. 2020. Unpacking the social media bot: A typology to guide research and policy. *Policy & Internet,* 12**,** 225-248.

GROVER, V. & LYYTINEN, K. 2015. New state of play in information systems research. *MIS quarterly,* 39**,** 271-296.

GURBAXANI, V. & DUNKLE, D. 2019a. Gearing up for successful digital transformation. *MIS Q. Executive,* 18**,** 6.

GURBAXANI, V. & DUNKLE, D. 2019b. Gearing up for successful digital transformation. *MIS Quarterly Executive,* 18.

HALL, M., FRANK, E., HOLMES, G., PFAHRINGER, B., REUTEMANN, P. & WITTEN, I. H. 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter,* 11**,** 10-18.

HARTL, E. & HESS, T. 2017. The role of cultural values for digital transformation: Insights from a Delphi study.

HAYES, A. F. & MATTHES, J. 2009. Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior research methods,* 41**,** 924-936.

HELFAT, C. E. 2007. Stylized facts, empirical research and theory development in management. Sage Publications Sage UK: London, England.

HENFRIDSSON, Y. Y. & LYYTINEN, K. 2010. The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. *Information Systems Research,* 21**,** 724-735.

HESS, T., MATT, C., BENLIAN, A. & WIESBÖCK, F. 2016. Options for formulating a digital transformation strategy. *MIS Quarterly Executive,* 15.

HIBBELN, M., JENKINS, J. L., SCHNEIDER, C., VALACICH, J. S. & WEINMANN, M. 2017. How is your user feeling? Inferring emotion through human–computer interaction devices. *Mis Quarterly,* 41**,** 1-22.

HITT, M. A., BEAMISH, P. W., JACKSON, S. E. & MATHIEU, J. E. 2007. Building theoretical and empirical bridges across levels: Multilevel research in management. *Academy of Management journal,* 50**,** 1385-1399.

HOBBS, J. R., STICKEL, M. E., APPELT, D. E. & MARTIN, P. 1993. Interpretation as abduction. *Artificial intelligence,* 63**,** 69-142.

HOFMANN, P., SAMP, C. & URBACH, N. 2020. Robotic process automation. *Electronic Markets,* 30**,** 99-106.

HSSINA, B., MERBOUHA, A., EZZIKOURI, H. & ERRITALI, M. 2014. A comparative study of decision tree ID3 and C4. 5. *International Journal of Advanced Computer Science and Applications,* 4**,** 13-19.

IM, K. S., DOW, K. E. & GROVER, V. 2001. A reexamination of IT investment and the market value of the firm—An event study methodology. *Information systems research,* 12**,** 103-117.

JARVIS, C. B., MACKENZIE, S. B. & PODSAKOFF, P. M. 2003. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of consumer research,* 30**,** 199-218.

JOVANOVIĆ, S. Z., ĐURIĆ, J. S. & ŠIBALIJA, T. V. 2018. Robotic process automation: overview and opportunities. *International Journal Advanced Quality,* 46**,** 34-39.

KAJROLKAR, A., PAWAR, S., PARALIKAR, P. & BHAGAT, N. Customer order processing using robotic process automation. 2021 International Conference on Communication information and Computing Technology (ICCICT), 2021. IEEE, 1-4.

KARHADE, P., SHAW, M. J. & SUBRAMANYAM, R. 2015. Patterns in information systems portfolio prioritization. *Mis Quarterly,* 39**,** 413-434.

KARIMI, J. & WALTER, Z. 2015. The role of dynamic capabilities in responding to digital disruption: A factor-based study of the newspaper industry. *Journal of Management Information Systems,* 32**,** 39-81.

KATHURIA, A., KARHADE, P. P. & KONSYNSKI, B. R. 2020. In the realm of hungry ghosts: multi-level theory for supplier participation on digital platforms. *Journal of Management Information Systems,* 37**,** 396-430.

KEDZIORA, D., LEIVONEN, A., PIOTROWICZ, W. & ÖÖRNI, A. 2021. Robotic Process Automation (RPA) Implementation Drivers: Evidence of Selected Nordic Companies. *Issues in Information Systems.*

KEDZIORA, D. & PENTTINEN, E. 2021. Governance models for robotic process automation: The case of Nordea Bank. *Journal of Information Technology Teaching Cases,* 11**,** 20-29.

KHANDAY, A. M. U. D., RABANI, S. T., KHAN, Q. R., ROUF, N. & MOHI UD DIN, M. 2020. Machine learning based approaches for detecting COVID-19 using clinical text data. *International Journal of Information Technology,* 12**,** 731-739.

KOHLI, R. & DEVARAJ, S. 2003. Measuring information technology payoff: A meta-analysis of structural variables in firm-level empirical research. *Information systems research,* 14**,** 127-145.

KONSYNSKI, B. R. & SVIOKLA, J. 1993. Cognitive reapportionment: Rethinking the location of judgement in managerial decision making. *Harvard Business School Discussion Draft, Harvard University.*

KOTSUKA, Y., MURAKAMI, R. & ARIMA, S. Development and Evaluation of Chat-bot Systems: The applications of citizen service procedures. 2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI), 2019. IEEE, 290-295.

LACITY, M. & WILLCOCKS, L. 2021. Becoming strategic with intelligent automation. *MIS Quarterly Executive,* 20**,** 1-14.

LACITY, M., WILLCOCKS, L. & GOZMAN, D. 2021. Influencing information systems practice: The action principles approach applied to robotic process and cognitive automation. *Journal of Information Technology,* 36**,** 216-240.

LAGNA, A. & RAVISHANKAR, M. 2022. Making the world a better place with fintech research. *Information Systems Journal,* 32**,** 61-102.

LANGLEY, A., MINTZBERG, H., PITCHER, P., POSADA, E. & SAINT-MACARY, J. 1995. Opening up decision making: The view from the black stool. *organization Science,* 6**,** 260-279.

LASSO-RODRIGUEZ, G. & WINKLER, K. 2020. Hyperautomation to fulfil jobs rather than executing tasks: the BPM manager robot vs human case. *Romanian Journal of Information Technology and Automatic Control,* 30**,** 7-22.

LIN, Y.-K., CHEN, H., BROWN, R. A., LI, S.-H. & YANG, H.-J. 2017. Healthcare Predictive Analytics for Risk Profiling in Chronic Care. *Mis Quarterly,* 41**,** 473-496.

LUO, Y., LIANG, P., WANG, C., SHAHIN, M. & ZHAN, J. Characteristics and challenges of low-code development: the practitioners' perspective. Proceedings of the 15th ACM/IEEE international symposium on empirical software engineering and measurement (ESEM), 2021. 1-11.

LYYTINEN, K., NICKERSON, J. V. & KING, J. L. 2021. Metahuman systems= humans+ machines that learn. *Journal of Information Technology,* 36**,** 427-445.

MACCALLUM, R. C. & BROWNE, M. W. 1993. The use of causal indicators in covariance structure models: some practical issues. *Psychological bulletin,* 114**,** 533.

MACKENZIE, S. B., PODSAKOFF, P. M. & JARVIS, C. B. 2005. The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. *Journal of applied psychology,* 90**,** 710.

MANDVIWALLA, M. & FLANAGAN, R. 2021. Small business digital transformation in the context of the pandemic. *European Journal of Information Systems,* 30**,** 359-375.

MARCH, J. G. 1994. *Primer on decision making: How decisions happen*, Simon and Schuster.

MARKUS, M. L., MAJCHRZAK, A. & GASSER, L. 2002. A design theory for systems that support emergent knowledge processes. *MIS quarterly***,** 179-212.

MARUPING, L. & MAGNI, M. 2012. What's the weather like. *The Effect of Team Learning Climate*.

MAYER, A.-S., STRICH, F. & FIEDLER, M. 2020. Unintended Consequences of Introducing AI Systems for Decision Making. *MIS Quarterly Executive,* 19.

MENDLING, J., DECKER, G., HULL, R., REIJERS, H. A. & WEBER, I. 2018. How do machine learning, robotic process automation, and blockchains affect the human factor in business process management? *Communications of the Association for Information Systems,* 43**,** 19.

MENDLING, J., PENTLAND, B. T. & RECKER, J. 2020. Building a complementary agenda for business process management and digital innovation. Taylor & Francis.

MIERS, D., KERREMANS, M., RAY, S. & TORNBOHM, C. 2019. Magic quadrant for robotic process automation software. *Stamford, CT: Gartner.*

MITHAS, S., XUE, L., HUANG, N. & BURTON-JONES, A. 2022. Editor's Comments: Causality Meets Diversity in Information Systems Research. *Management Information Systems Quarterly,* 46**,** iii-xviii.

MOHANTY, S. & VYAS, S. 2018. *How to compete in the age of artificial intelligence: Implementing a collaborative human-machine strategy for your business*, Apress.

NAQVI, A. & MUNOZ, J. M. 2020. *Handbook of Artificial Intelligence and Robotic Process Automation: Policy and Government Applications*, Anthem Press.

OSEI-BRYSON, K. M. & NGWENYAMA, O. 2011. Using decision tree modelling to support Peircian abduction in IS research: a systematic approach for generating and evaluating hypotheses for systematic theory development. *Information Systems Journal,* 21**,** 407-440.

OSHRI, I. & PLUGGE, A. 2022. Introducing RPA and automation in the financial sector: Lessons from KAS Bank. *Journal of Information Technology Teaching Cases,* 12**,** 88-95.

OSMUNDSEN, K., IDEN, J. & BYGSTAD, B. 2019. Organizing robotic process automation: balancing loose and tight coupling.

OTIM, S. & GROVER, V. 2012. Resolving uncertainty and creating value from the exercise of e-commerce investment options. *Information Systems Journal,* 22**,** 261-287.

ØVRELID, E. & BYGSTAD, B. 2019. The role of discourse in transforming digital infrastructures. *Journal of Information Technology,* 34**,** 221-242.

PARÉ, G., TRUDEL, M.-C., JAANA, M. & KITSIOU, S. 2015. Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management,* 52**,** 183-199.

PARUCHURI, S., PERRY-SMITH, J. E., CHATTOPADHYAY, P. & SHAW, J. D. 2018. New ways of seeing: Pitfalls and opportunities in multilevel research. Academy of Management Briarcliff Manor, NY.

PENTLAND, B. T. 1999. Building process theory with narrative: From description to explanation. *Academy of management Review,* 24**,** 711-724.

PENTTINEN, E., KASSLIN, H. & ASATIANI, A. How to choose between robotic process automation and back-end system automation? European Conference on Information Systems 2018, 2018.

PETTER, S., STRAUB, D. & RAI, A. 2007. Specifying formative constructs in information systems research. *MIS quarterly***,** 623-656.

PLATTFAUT, R. 2019. Robotic Process Automation–Process Optimization on Steroids?

POMEROL, J.-C., BRÉZILLON, P. & PASQUIER, L. 2002. Operational knowledge representation for practical decision-making. *Journal of Management Information Systems,* 18**,** 101-115.

QUINLAN, J. R. 1986a. Induction of decision trees. *Machine learning,* 1**,** 81-106.

QUINLAN, J. R. 1986b. Induction of decision trees. *Machine learning,* 1**,** 81-106.

QUINLAN, J. R. 1990. Decision trees and decision-making. *IEEE Transactions on Systems, Man, and Cybernetics,* 20**,** 339-346.

QUINLAN, J. R. 2014. *C4. 5: programs for machine learning*, Elsevier.

RANA, N. P., CHATTERJEE, S., DWIVEDI, Y. K. & AKTER, S. 2022. Understanding dark side of artificial intelligence (AI) integrated business analytics: assessing firm's operational inefficiency and competitiveness. *European Journal of Information Systems,* 31**,** 364-387.

REICHERTZ, J. 2007. Abduction: The logic of discovery of grounded theory. *The SAGE handbook of grounded theory***,** 214-228.

RIEMER, K. & PETER, S. 2020. The robo-apocalypse plays out in the quality, not in the quantity of work. *Journal of Information Technology,* 35**,** 310-315.

SABHERWAL, R. & CHAN, Y. E. 2001. Alignment between business and IS strategies: A study of prospectors, analyzers, and defenders. *Information systems research,* 12**,** 11-33.

SÆTRE, A. S. & VAN DE VEN, A. 2021. Generating theory by abduction. *Academy of Management Review,* 46**,** 684-701.

SANDBERG, J., MATHIASSEN, L. & NAPIER, N. 2014. Digital options theory for IT capability investment. *Journal of the Association for Information Systems,* 15**,** 1.

SANTOS, F., PEREIRA, R. & VASCONCELOS, J. B. 2019. Toward robotic process automation implementation: an end-to-end perspective. *Business Process Management Journal,* 26**,** 405-420.

SCHANKE, S., BURTCH, G. & RAY, G. 2021. Estimating the impact of "humanizing" customer service chatbots. *Information Systems Research,* 32**,** 736-751.

SEILONEN, I., PIRTTIOJA, T., APPELQVIST, P., HALME, A. & KOSKINEN, K. Distributed planning agents for intelligent process automation. Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation. Computational Intelligence in Robotics and Automation for the New Millennium (Cat. No. 03EX694), 2003. IEEE, 614-619.

SIA, S. K., SOH, C. & WEILL, P. 2016. How DBS Bank Pursued a Digital Business Strategy. *MIS Quarterly Executive,* 15.

SOLUK, J. & KAMMERLANDER, N. 2021. Digital transformation in family-owned Mittelstand firms: A dynamic capabilities perspective. *European Journal of Information Systems,* 30**,** 676-711.

SOMEH, I., WIXOM, B. H., BEATH, C. M. & ZUTAVERN, A. 2022. Building an Artificial Intelligence Explanation Capability. *MIS Quarterly Executive,* 21**,** 5.

SPANGLER, W. E., MAY, J. H. & VARGAS, L. G. 1999. Choosing data-mining methods for multiple classification: representational and performance measurement implications for decision support. *Journal of Management Information Systems,* 16**,** 37-62.

STRICH, F., MAYER, A.-S. & FIEDLER, M. 2021. What do I do in a world of Artificial Intelligence? Investigating the impact of substitutive decision-making AI systems on employees' professional role identity. *Journal of the Association for Information Systems,* 22**,** 9.

STURM, T., GERLACH, J. P., PUMPLUN, L., MESBAH, N., PETERS, F., TAUCHERT, C., NAN, N. & BUXMANN, P. 2021. Coordinating Human and Machine Learning for Effective Organizational Learning. *MIS Quarterly,* 45.

SUH, A., SHIN, K.-S., AHUJA, M. & KIM, M. S. 2011. The influence of virtuality on social networks within and across work groups: A multilevel approach. *Journal of Management Information Systems,* 28**,** 351-386.

SYED, R., SURIADI, S., ADAMS, M., BANDARA, W., LEEMANS, S. J., OUYANG, C., TER HOFSTEDE, A. H., VAN DE WEERD, I., WYNN, M. T. & REIJERS, H. A. 2020. Robotic process automation: contemporary themes and challenges. *Computers in Industry,* 115**,** 103162.

TAN, F. T., ONDRUS, J., TAN, B. & OH, J. 2020. Digital transformation of business ecosystems: Evidence from the Korean pop industry. *Information Systems Journal,* 30**,** 866-898.

TAULLI, T. 2020. The robotic process automation handbook. *The Robotic Process Automation Handbook. https://doi. org/10.1007/978-1-4842-5729-6.*

TEMPLIER, M. & PARE, G. 2018. Transparency in literature reviews: an assessment of reporting practices across review types and genres in top IS journals. *European Journal of Information Systems,* 27**,** 503-550.

TESSMER, A. C., SHAW, M. J. & GENTRY, J. A. 1993. Inductive learning for international financial analysis: a layered approach. *Journal of Management Information Systems,* 9**,** 17-36.

TESSMER, M. 1994. Formative evaluation alternatives. *Performance Improvement Quarterly,* 7**,** 3-18.

THEKKETHIL, M. S., SHUKLA, V. K., BEENA, F. & CHOPRA, A. Robotic process automation in banking and finance sector for loan processing and fraud detection. 2021

9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), 2021. IEEE, 1-6.

TILSON, D., LYYTINEN, K. & SØRENSEN, C. 2010. Research commentary—Digital infrastructures: The missing IS research agenda. *Information systems research,* 21**,** 748-759.

TIMBADIA, D. H., SHAH, P. J., SUDHANVAN, S. & AGRAWAL, S. Robotic process automation through advance process analysis model. 2020 International Conference on Inventive Computation Technologies (ICICT), 2020. IEEE, 953-959.

TORKHANI, R., LAVAL, J., MALEK, H. & MOALLA, N. 2019. Collaborative information system toward a robotic process automation solution. *Enterprise Interoperability VIII.* Springer.

VIAL, G. 2021. Understanding digital transformation: A review and a research agenda. *Managing Digital Transformation***,** 13-66.

VIALE, L. & ZOUARI, D. Impact of digitalization on procurement: the case of robotic process automation. Supply Chain Forum: An International Journal, 2020. Taylor & Francis, 185-195.

VIEHHAUSER, J. Is robotic process automation becoming intelligent? Early evidence of influences of artificial intelligence on robotic process automation. Business Process Management: Blockchain and Robotic Process Automation Forum: BPM 2020 Blockchain and RPA Forum, Seville, Spain, September 13–18, 2020, Proceedings 18, 2020. Springer, 101-115.

VOM BROCKE, J., SIMONS, A., RIEMER, K., NIEHAVES, B., PLATTFAUT, R. & CLEVEN, A. 2015. Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the association for information systems,* 37**,** 9.

WANG, G., CHEN, G., ZHAO, H., ZHANG, F., YANG, S. & LU, T. 2021. LEVERAGING MULTISOURCE HETEROGENEOUS DATA FOR FINANCIAL RISK PREDICTION: A NOVEL HYBRID-STRATEGY-BASED SELF-ADAPTIVE METHOD. *MIS Quarterly,* 45.

WATSON, R. T. & WEBSTER, J. 2020. Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0. *Journal of Decision Systems,* 29**,** 129-147.

WEICK, K. E. 1995. What theory is not, theorizing is. *Administrative science quarterly,* 40**,** 385-390.

WEISS, S. M. & INDURKHYA, N. Decision tree pruning: biased or optimal? AAAI, 1994. 626-632.

WESSEL, L., BAIYERE, A., OLOGEANU-TADDEI, R., CHA, J. & BLEGIND-JENSEN, T. 2021. Unpacking the difference between digital transformation and IT-enabled organizational transformation. *Journal of the Association for Information Systems,* 22**,** 102-129.

WEWERKA, J. & REICHERT, M. Towards quantifying the effects of robotic process automation. 2020 IEEE 24th International Enterprise Distributed Object Computing Workshop (EDOCW), 2020. IEEE, 11-19.

WILLCOCKS, L., LACITY, M. & CRAIG, A. 2017. Robotic process automation: strategic transformation lever for global business services? *Journal of Information Technology Teaching Cases,* 7**,** 17-28.

WIMELIUS, H., MATHIASSEN, L., HOLMSTRÖM, J. & KEIL, M. 2021. A paradoxical perspective on technology renewal in digital transformation. *Information systems journal,* 31**,** 198-225.

XIE, K. & LEE, Y.-J. 2015. Social media and brand purchase: Quantifying the effects of exposures to earned and owned social media activities in a two-stage decision making model. *Journal of Management Information Systems,* 32**,** 204-238.

YAHAV, I., SHMUELI, G. & MANI, D. 2016. A tree-based approach for addressing self-selection in impact studies with big data. *MIS Quarterly,* 40**,** 819-848.

ZAPADKA, P., HANELT, A. & FIRK, S. 2022. Digital at the edge–antecedents and performance effects of boundary resource deployment. *The Journal of Strategic Information Systems,* 31**,** 101708.

12 APPENDIX A

Table 35: Summary of Prior IS Research of Digital Transformation.

Study	Theory Base	DV	IV(Moderator/Mediator)	Method	Key Arguments & Findings	Limitation
(Baiyere et al., 2020)	BPM needs/logics in the context of DT	Dynamics of BPM in the context of DT	Light touch processes infrastructural flexibility mindful actors	Ethnographic study	In the context of digital transformation, the needs of BPM change because of ever-changing DT. Light touch processes, infrastructural flexibility, and mindfulness of actors should stimulate the imaginations of BPM scholars and practitioners alike.	Just studied on one company. Diversity is low. Very generic. Empirically not proven. Theory is not generated based on the data.
(Mandviwalla and Flanagan, 2021)	Action design research (ADR)	Small business firm value through DT	Engagement Selling Delivery Process/new Models	Case study	DT is welcome and can generate values in small businesses. Digital technologies, especially platforms that target small. Businesses have matured to a level that can accelerate the transformation. Small businesses should focus initially on digital channel basics such as engaging, selling, delivering, and over time expanding to explore new digital business models.	Results are limited to microbusinesses. Diversity is low. Very generic. Empirically not proven. Theory is not generated based on the data.
(Soluk and Kammerlander, 2021)	Dynamic capabilities	Achieve digital transformation	Strategic decision-making Information management Continuous renewal Employee learnability Strategic partnerships Brand management	Case study	Challenges and opportunities in DT will continue. The antecedents, processes, and implications of firms' digital transformation for information systems and management research are still under-researched.	Managerial implications for DT are limited to family-owned firms. Empirically not proven. Theory is not generated based on the data

					Analyzes the processual dimension of digital transformation in family-owned Mittelstand firms. Triggers, enablers, and barriers in the digital transformation process, and the role of dynamic capabilities in achieving digital transformation.	
(Wimelius et al., 2021)	Paradoxical tensions for technology renewal	Technology renewal	Paradoxical tensions, i.e., (established vs renewed technology usage), deliberate vs emergent renewal practices, inner vs outer renewal contexts.	Longitudinal case study	Technology renewal of digital and infrastructure platforms is essential to achieve strategic goals. Technology renewal involves paradoxical tensions between established and renewed technology usage, deliberate and emergent renewal practices, and inner and outer renewal contexts. Organizations respond differently to the paradoxical tensions. Reinforcing a virtuous cycle and increasing the likelihood of renewal success requires persistent patterns of integrating and splitting responses. In contrast, persistent patterns of pretending and avoiding responses will reinforce a vicious cycle and increase the likelihood of renewal failure.	Qualitative explanation. Diversity is low. Very generic. Empirically not proven. Theory is not generated based on the data.
(Tan et al., 2020)	Driving boundary practices to digitally transform business ecosystems. Democratization as	Digital transformation of business ecosystem	Boundary practices	Interpretive case study	Insights into the process of digital transformation of business ecosystems. Resources can be combined to effectively	Problem of transferability or generalizability. Limited context. Limited data.

	boundary practice				develop business ecosystems.	Empirically not proven. Theory is not generated based on the data.
(Chanias et al., 2019)	Integrated process/activity model showing how pre-digital organizations can develop a digital transformation strategy	Realized digital transformation strategy	Organizational strategy Episodes of digital strategy DTS practices	Interpretive in-depth case study	Human agents had to continually work at transforming the social structure. Digital strategy making where a DTS must be continually reinvented. Digital strategy has distinguished characteristics compared with IS strategy. Digital transformation is business-centric and customer-oriented in its perspective. All parts of the organization are affected by changes resulting from a DTS. DTS is developed by different stakeholders within the organization.	Just one case study. Problem of transferability or generalizability. Limited context. Limited data. Empirically not proven. Theory is not generated based on the data.
(Zapadka et al., 2022)	Boundary resource deployment	Beneficial boundary resources	Digital knowledge Digital complementors Market power	Empirical study on longitudinal sample. General estimation of equations (GEE) regression	Firms that rank high in digital knowledge tend to deploy boundary resources. Existence of digital complementors in the field is positively associated with boundary resource deployments.	Restricted sample. Geographical limitation. Limited context. Limited data. Empirically not proven. Theory is not generated based on the data.
(Wessel et al., 2021)	Grounded theory	Differences between DT and ITOT	Transformation activities Transformation outcome	Longitudinal case study	Conceptual differences between DT and ITOT. DT changes the identity of the firm. ITOT reinforces existing organizational identity.	Theory is not tested. Limited to few case studies. Theory is not generated based on the data.
(Sandberg et al., 2014)	General options theory	Digital options	Connectivity Uncertainty Equivocality Context appreciation Characterization Information requirements analysis	Case study	Consider technology innovation in relation to sociotechnical changes.	Only limited to conceptual foundation. Empirically not proven. Theory is not generated

			Digital options recognition.		Should consider technical and rational decisions and cultural, social, and cognitive forces. Identifying suitable processes for IT investment is a key activity in the context appreciation activity.	based on the data.
(Gurbaxani and Dunkle, 2019b)	Six-dimension framework for executives	Successful digital transformation	Strategic vision Culture of innovation Expertise and intellectual property Dimension: know-how and intellectual property Digital capability Strategic alignment Technology assets	Exemplar case studies	Provides framework for executives to assess their company's progress on six dimensions critical to successful digital transformation. Benchmarking one's company with others in our database—either within a sector or against companies that are in the same state of progress toward digital transformation. Helps diagnose gaps in a company's capabilities.	Basic research. Empirically not proven. Theory is not generated based on the data.
(Hess et al., 2016)	Conceptual framework for formulating a digital transformation strategy and key dimensions	Dimensions of digital transformation	Use of technologies Value creation Structural changes Financial aspects	Case study of three companies	Conceptual framework for formulating a digital transformation strategy. What are the right questions to ask? Provide managers with a comprehensive and structured approach to digital transformation.	Basic research. Empirically not proven. Theory is not generated based on the data.
(Karimi and Walter, 2015)	Dynamic capabilities	Response to disruptive innovation	Dynamic capabilities	Case study	Clarifies the role of first-order dynamic capabilities in responding to digital disruption. Helps building digital platform capabilities, and for reinventing their core functions to accelerate digitization.	Restricted sample. Geographical limitation. Limited context. Limited data. Empirically not proven. Theory is not generated based on the data.
(Kohli and Devaraj, 2003)	A Framework for the structural categories influencing IT payoff	Firm performance/profitability	Type of domain Sample size Data source	Meta-analysis Logistic regression Discriminant analysis	A framework for the structural categories influencing IT payoff	Empirically proven. Meta-analysis includes studies from the

						information systems discipline. Real implementation of data is not considered.
(Øvrelid and Bygstad, 2019)	Foucault's theory of discourse	Digital infrastructure transformation	Contextual factors Causal mechanisms	Case study	Framework to understand the role of discursive formations in digital transformation. Propose a set of configurations to explain how contextual factors and causal mechanisms contingently lead to the transformation of a digital infrastructure.	Experimental. Theoretical. Theory is not generated based on the data.
(Datta et al., 2020)	Digital transformation challenges	Digital transformation success	Sociocultural disruption Digital literacy Bureaucratic friction	Case study	Offers digital transformation recommendations, generalizable across any global democracy	Narrowed focus. No quantitative proof. Theory is not generated based on the data.
(Chaimankong et al., 2021)	To explore, understand, analyze, and summarize the impacts of the COVID-19 pandemic on digital banking services	Digital transformation impact or customer retention	Consumer behavior Consumer trends Service innovation Customer engagement	Semistructured interviews	Customers want better experience and to manage their finances conveniently from any location.	Context sensitive. Geographical limitation. Limited context.
(Sabherwal and Chan, 2001)	Miles and Snow's popular classification of defender, analyzer, and prospector business strategies. STROBE framework.	Business success or firm performance	Alignment between business and IS strategy	Two multiresponde nt surveys Empirical methods	Alignment influences overall business success in prospectors and analyzers but not in defenders. Aligning the IS strategy with the business strategy may not be as universal as previously believed.	Simplification. Applicability to other industries. No objective measures. Not based on real data. Theory is not generated based on the data.
(Im et al., 2001)	Changes in the market value of the firm	Effectiveness of IT investments	Price reaction Volume reaction Industry effect Size effect Time lag effect	The event study methodology Statistical methods	There is no price reaction for larger firms and a positive price reaction for smaller firms. There is an increase in both price and volume reaction over time. Both industry and size effects become stronger over time	All potential confounding variables were not considered. Only related to stock prices. Theory is not generated based on the data.

					IT spending is of value to the firm.	
(Sia et al., 2016)	Pursuing a digital business strategy	Digital strategy success factors	Structure process technology people	Case study	Key capabilities that an organization needs to build so it can pursue a digital business strategy. There is greater urgency to "rewire" or transform traditional enterprises so they can accommodate digital innovation. Important questions to ask.	Case study restricted to the banking domain. Narrowed focus. No quantitative proof. Theory is not generated based on the data.

13 APPENDIX B

Table 36: Summary of Prior IS Research of Artificial Intelligence.

Study	Theory Base	DV	IV(Moderator/Mediator)	Method	Key Arguments & Findings	Limitation
(Benbya et al., 2021)	Tensions of AI for information systems	AI tensions	Substitution of jobs vs. tasks Automation vs. Augmentation Humanlike vs. machinelike conversations Human vs. artificial emotion intelligence Machine rationality vs. human judgment Human vs. machine bias Decision accountability humans vs. machines	Literature review and case study	Differentiated effects that AI brings about and the implication for future IS research.	Qualitative case study approaches. Empirically proven. No quantitative proof. Theory is not generated based on the data.
(Strich et al., 2021)	Mechanisms through which employees strengthen and protect their professional role identity.	Role identity before and after AI	Foresighted consulting enhanced consulting services, data manipulation, self-elevation, responsibility transfer, illustration of consultation, reassurance	Case study	Shedding light on how a substitutive decision-making AI system affects employees' professional role identity. Revealing different mechanisms utilized by the two consultant groups to respond to the changes in their professional role identities. Highlighting the boundary conditions resulting from introducing a substitutive decision-making AI system. Contribution to the empirical literature on AI and employees.	Focusses on single point of time. Single case study. No quantitative proof. Theory is not generated based on the data.
(Riemer and Peter, 2020)	Willcocks' analysis of the automation and future of work	Effect on work life quality	Task complexity Skills development Job control Work intensity Nature of work Arrangements Perceived job security. Perceived job security	Literature review and Wilcocks' analysis	Automation and the future of work must include the qualitative changes automation will bring to work and workplaces. Suggest that aspects of job design and employee experience should become part of new automation initiatives so that automation does not invariably result in unintended outcomes for work life quality. Most countries lack the capability to keep track of changes in the qualitative aspects of work in their economies.	No quantitative proof. Theory is not generated based on the data.
(Khanday et al., 2020)	Help of various AI tools	Diagnosis of disease	Image data Textual data 24 Attributes	Data analysis Classification	Revealed that logistic regression and multinomial naive	Quantitatively proven.

				Logistic regression Multinomial naıve Bayes Support vector machine Decision trees Bagging AdaBoost Random forest Stochastic gradient boosting	Bayesian classifier gives excellent results by having 94% precision, 96% recall, 95% f1 score, and accuracy 96.2% Various other machine-learning algorithms that showed better results were random forest, stochastic gradient boosting, decision trees and boosting. The efficiency of models can be improved by increasing the amount of data.	Needs more data. Theory is not explained only results have been explained.
(Fügener et al., 2021)	Attitude toward AI Delegation Complementarity on the instance level Role of feedback	Accuracy	AI alone Humans alone Delegation Inversion	Experimental design Descriptive statistics	Demonstrates that humans and AI can work together. If AI would be responsible to delegate to humans, the resulting performance was higher than that of the AI alone. Inversion might also improve human work perspectives. Humans making more arbitrary delegation decisions when dealing with difficult tasks, which worsens their overall performance.	Not generalizable. Restricted to nonspecialized situations. Theory is not generated based on the data across situations.
(Berente et al., 2021b)	Synthesizes the insights on managing AI	Facets of AI	Autonomy Learning Inscrutability	Exemplar case studies	Reflects about how our own norms, processes, outputs, and "ground truth" may be challenged in terms of autonomy, learning, and inscrutability	Narrowed focus. No quantitative proof. Theory is not generated based on the data.
(Schanke et al., 2021)	Anthropomorphism of AI-enabled automated customer service	Transaction outcomes	Social presence Communication delays Humor	Field experiment Descriptive statistics	Anthropomorphism influences transaction conversion positively. Anthropomorphism, in our context, plays the most significant role in sensitive information disclosure. Augmenting AI-enabled autonomous agents with human-like social intelligence can increase their performance in customer service settings.	Narrowed focus. Theory is not generated based on the data.
(Someh et al., 2022)	Inductive grounded theory. Challenges and explains the ability dimensions of AI	AI explainability	Decision tracing Bias remediation Boundary setting Value formulation	Case study	Include and engage with the entire organization to Build AIX capability. Look beyond the AI team to assemble the required AI explanation expertise. Document current practices for decision tracing, bias remediation,	Narrowed focus. Theory is not generated based on the data.

					boundary setting, and value formulation.	
(Mayer et al., 2020)	Unintended consequences of introducing a decision-making AI system	Unintended consequences of AI	AI system	Case study	Highlights the potential benefits of substituting human decision-making with an AI system. Was confronted with several unintended consequences of introducing the AI system, for both frontline employees and the organization, which were not anticipated by senior management during the planning stages. Unintended consequences could threaten the intended organizational goals of introducing an AI system. Provided recommendations for managers who intend to implement or have already implemented AI systems in their organization.	Narrowed focus. Theory is not generated based on the data.
(Rana et al., 2022)	Dynamic capability view and contingency theory Resource-based view Unintended consequences of AI-integrated business analytics (AI-BA) influence a firm's overall competitive advantage	Adverse firm performance	Opacity Operational inefficiency Contingency planning Competitive disadvantage.	Interviews Descriptive statistics	AI-integrated BA solution might delve into several misplaced assumptions where the potential dangers might be introduced by AI in the firm settings. Effective administration of AI governance in a firm brings sustenance toward competitiveness of that firm. Ineffective AI governance would negatively influence the performance of the firm, and in that way, the firm would lose its competitiveness through operational inefficiency.	Narrow focus to service industries. Low sample sizes. Overlooked technical issues.

14 APPENDIX C

Table 37: Summary of Prior IS Research of Intelligence Process Automation.

Study	Theory Base	DV	IV(Moderator/Mediator)	Method	Key Arguments & Findings	Limitation
(Denagama Vitharanage et al., 2020)	Empirical studies on RPA benefits gained by organizations	RPA benefits	Accuracy Average handling time Process efficiency ROI Customer satisfaction	Exploratory case study	"Improvement in accuracy" was the most discussed anticipated benefit, while "improvement in customer service and customer satisfaction" was the least discussed anticipated benefit. Identified seven anticipated benefits and seven unanticipated benefits.	Single case study. Single process
(Plattfaut, 2019)	Identify key lessons learned in RPA	Lessons learned RPA	Test beyond technology Program communication IT and business commitment Prioritization	Case study	Extends this test from a pure technical one to a test also including regulatory and governance issues. Includes RPA in the overall process optimization program communication. Business needs to be behind the technology and committed to RPA introduction. Not overthink prioritization procedures. Organizations need to source lasting capabilities.	Single case study.
(Carden et al., 2019)	Resources and tools and techniques related to project execution	RPA outcomes	Cost Efficiency Cycle time or handling time	Case study	What are the resources, tools, and techniques related to project execution? Future issues and challenges related to robotics process automation, cognitive tools, and blockchain integration.	Single case study.
(Asatiani and Penttinen, 2016)	Challenges for RPA implementation	RPA introduction	Potential analysis Process assessment Business case	Case study	RPA business model is dependent on short- and long-term goals. Analyzes the market opportunity.	Single case study
(Lacity et al., 2021)	Guidelines for RPA action principles	RPA action principles	Strategy Sourcing Program management Process selection Tool selection Stakeholder buy-in	Case study	Workable approach that gains a great deal of contemporary information, providing insights into how the technologies function, how they are deployed, and with what results.	Single case study

(Kedziora and Penttinen, 2021)	Discussion about RPA journey	RPA governance	Exception handling software Integration Reusability Downtime Standardization IT complexity	Case study	Outlines several governance-related issues and decision points that must be addressed in connection with any deployment of robotic process automation. The key issues are related to the software's development and maintenance, robotic process automation governance, and IT infrastructure.	Sole case study
(Oshri and Plugge, 2022)	Journey to implement RPA solutions	RPA introduction	Process feasibility Service quality Customer satisfaction	Case study	Understanding what bots can and cannot do. Understanding the end-to-end business process. When the bot fails to complete a task.	Sole case study. Empirically not proven.
(Lyytinen et al., 2021)	Metahuman systems	Metahuman critical factors	Delegating Monitoring Cultivating Reflecting	Literature survey	Addresses issues of human goals and values in settings where metahuman systems evolve or are applied. Achieving benefits and avoiding problems will require better understanding of systems level learning.	Literature review. Empirically not proven.
(Lyytinen et al., 2021)	A Framework for explaining the behavior of black-box AI systems	AI explanation	Model Goals Training data Input data Output data Environment	Case study	Framework for explaining the behavior of black-box AI systems can facilitate the successful introduction of AI.	Sole case study Empirically not proven
(Mendling et al., 2020)	Orthogonal assumptions of process management and digital innovation	Convergent logic	Combined process design	Exemplar case studies	BPM and digital innovation belong together, like two sides of the same coin. BPM and digital innovation are complementary fields of inquiry that have much to learn from, and offer to, each other. Processes, technologies, and products are intertwined.	
(Bygstad and Øvrelid, 2020)	Investigates the alignment between process innovation and architectural alignment	Successful process innovation and digital infrastructure alignment	Lightweight IT Vendor boundary resources Message exchange	Case study	The careful deployment of lightweight IT in onsite configuration, loosely coupled from the infrastructure activities, allows for fast process innovation while leveraging the slow and nonlinear evolution of infrastructure.	

					Model to describe the interaction between lightweight IT and heavyweight for process innovation efforts to successfully interact and align with a large existing digital infrastructure.	

15 APPENDIX D – WEKA FOR DECISION TREE INDUCTION

Figure 13: Weka Tool for Decision Tree Induction

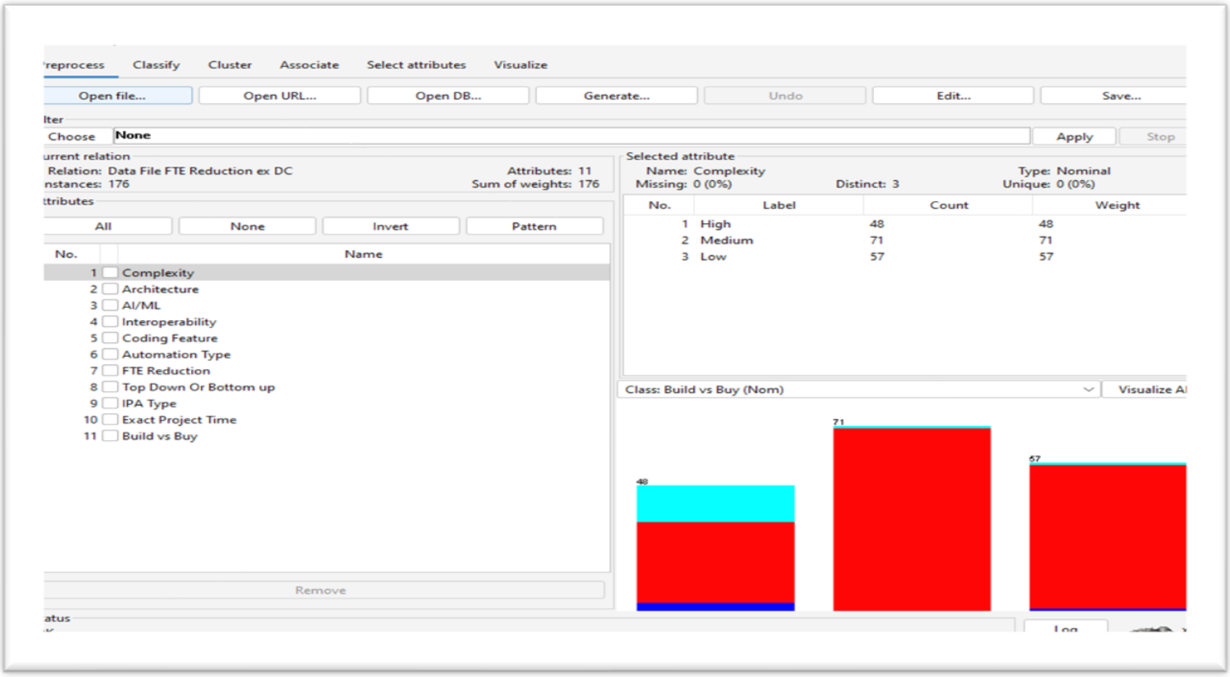

Figure 14: Data Preprocessing in Weka

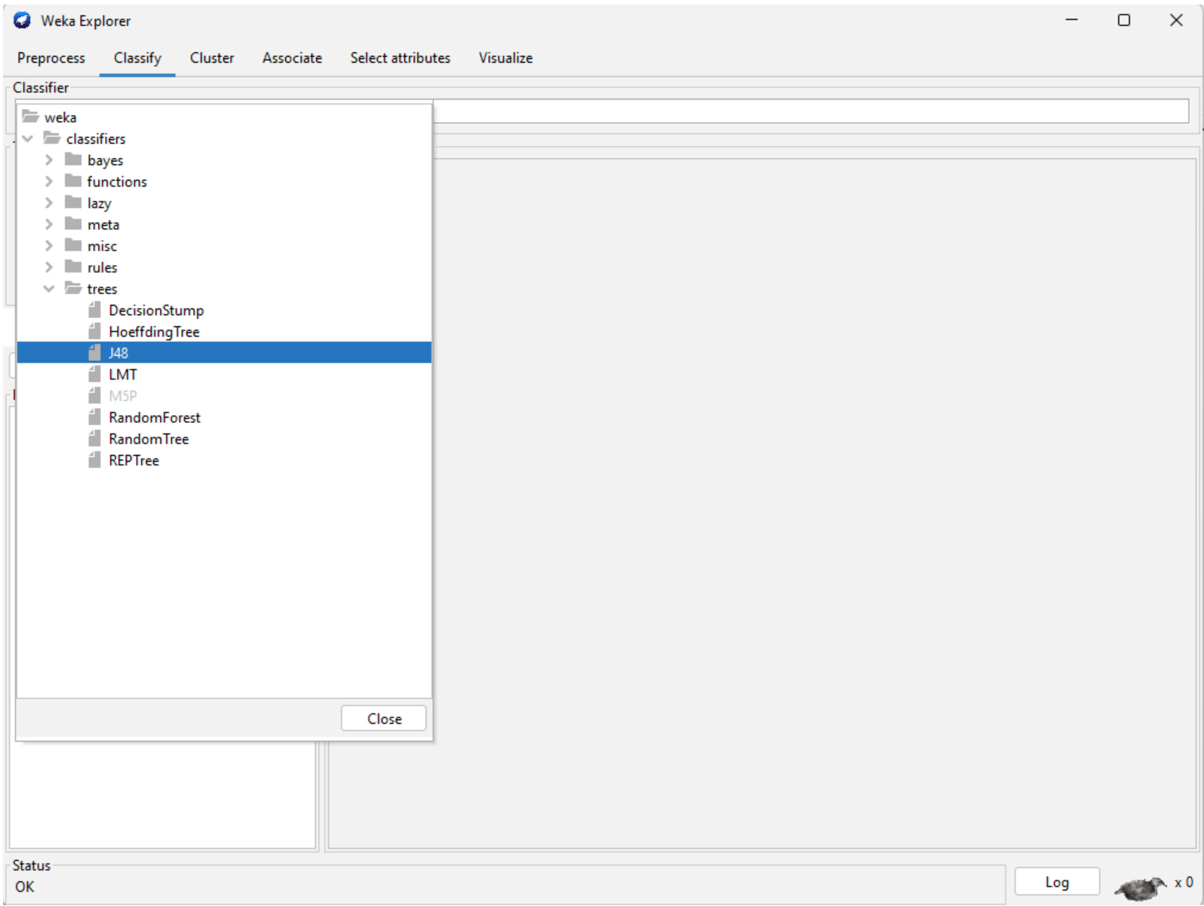

Figure 15: Data Classification Using C4.5 Decision Tree Induction

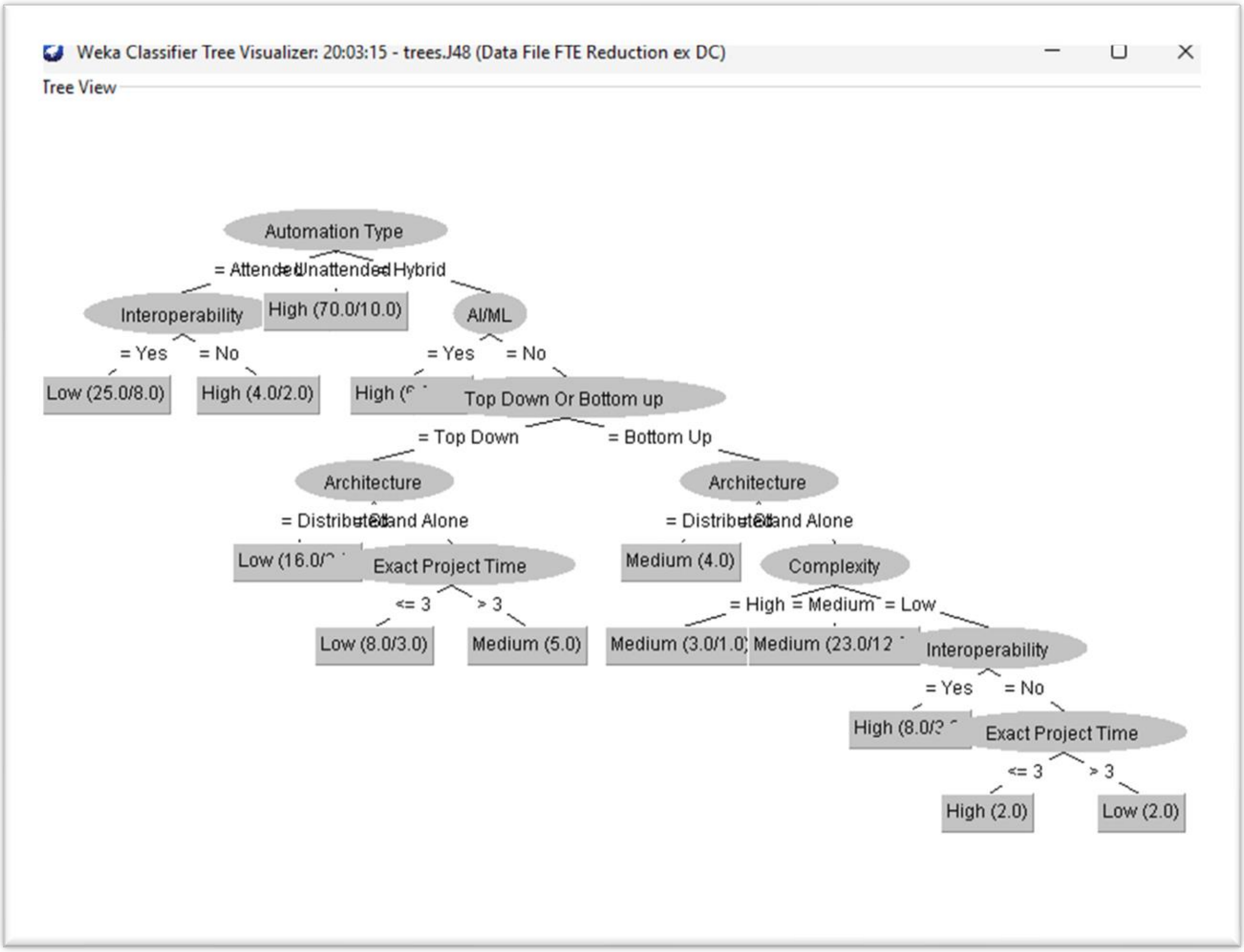

Figure 16: Representation of decision tree for FTE Reduction in Weka

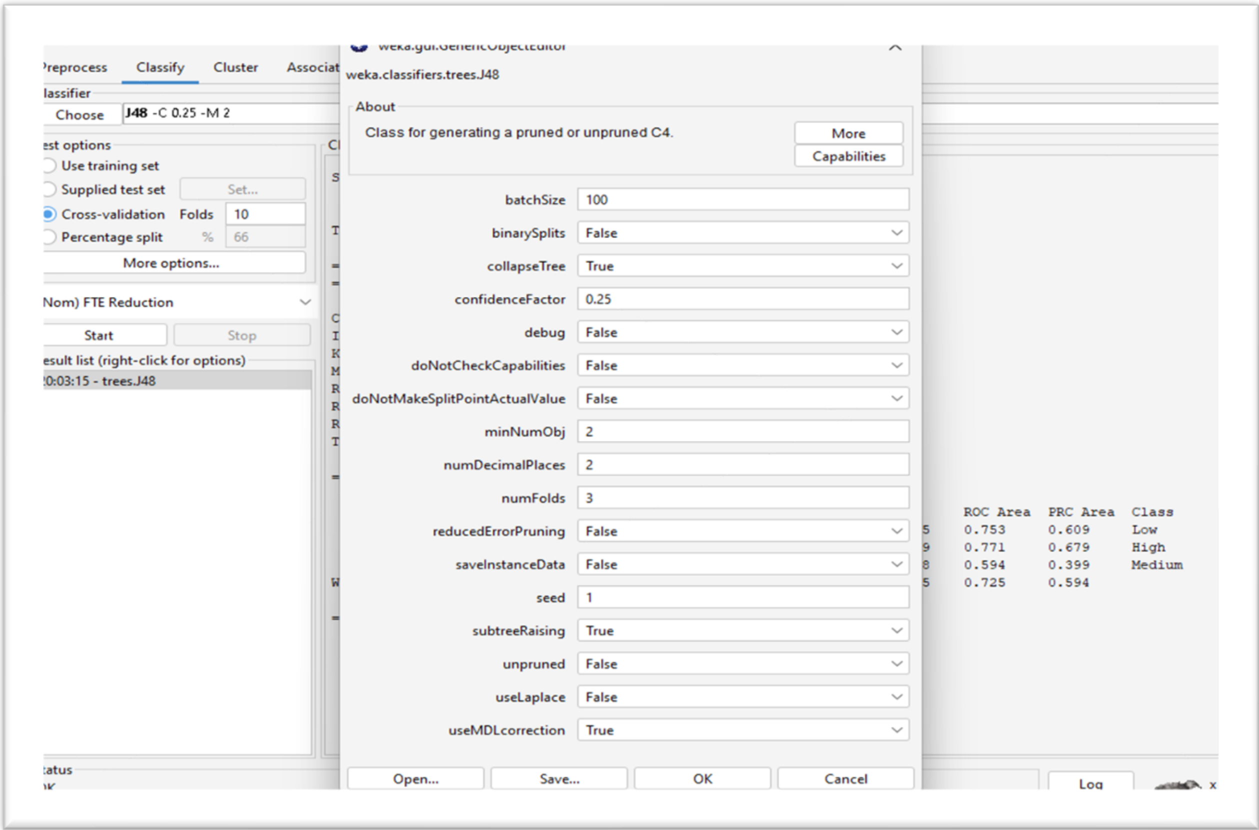

Figure 17: Decision Tree Pruning using Weka.