From 0 to 100 – AI Maturity and Business Transformation

Ben Schütte

Dissertation written under the supervision of professor

René Bohnsack

Dissertation submitted in partial fulfilment of requirements for the MSc in Business at the Universidade Católica Portuguesa, 04.01.2026.

Abstract

This thesis investigates how organizations measure the return on investment (ROI) of artificial intelligence (AI) initiatives and which practices and challenges shape this evaluation in practice by drawing on qualitative expert interviews, an extensive review of the academic literature, and an analysis of corporate reporting practices. It is shown that the measurement of artificial intelligence return on investment cannot be reduced to a single metric or evaluation logic, as value creation occurs across multiple organizational levels and includes financial outcomes as well as capability development, learning effects, and strategic positioning. Accordingly, heterogeneous measurement approaches are applied depending on the scope of artificial intelligence implementation and the type of anticipated value. Based on the empirical findings, a framework for artificial intelligence return on investment measurement is developed that differentiates between project-based and organization-wide initiatives as well as between soft and hard return orientations, resulting in four overarching measurement logics. These comprise adoption and fluency scorecards, use-case-specific financial evaluations based on counterfactuals, governance-related value indicators, and enterprise-level artificial intelligence infrastructure economics. Rather than offering a prescriptive solution, the framework provides an integrative lens that structures existing practices and clarifies the conditions under which different measurement approaches are applied. By doing so, the study contributes to the literature on artificial intelligence and strategic management by advancing the understanding of how organizations assess the value of artificial intelligence investments and by deriving practical implications for the design of measurement systems that capture both short-term performance and long-term strategic impact.

Keywords: Artificial Intelligence, Return on Investment, AI Value Measurement, Digital Transformation, Strategic Management

Title: From 0 to 100 – AI Maturity and Business Transformation

Author: Ben Schütte

Resumo

Esta dissertação analisa a forma como as organizações medem o retorno sobre o investimento (ROI) de iniciativas de inteligência artificial (IA), com base em entrevistas qualitativas com especialistas, numa revisão da literatura académica e numa análise de práticas de reporte corporativo. Demonstra-se que a medição do ROI em IA não pode ser reduzida a uma métrica única, uma vez que a criação de valor ocorre em diferentes níveis organizacionais e abrange resultados financeiros, desenvolvimento de capacidades, efeitos de aprendizagem e posicionamento estratégico. Em consequência, são adotadas abordagens de medição diferenciadas, consoante o âmbito da implementação da IA e o tipo de valor esperado. Com base nos resultados empíricos, é desenvolvido um enquadramento para a medição do ROI em IA que distingue entre iniciativas baseadas em projetos e iniciativas ao nível organizacional, bem como entre orientações de retorno suave e retorno duro. Dessa distinção resultam quatro lógicas gerais de medição, incluindo scorecards de adoção e fluência, avaliações financeiras específicas por caso de uso baseadas em contrafactuais e indicadores de valor associados à governação e à economia da infraestrutura de IA ao nível empresarial. Em vez de propor uma solução prescritiva, o enquadramento fornece uma perspetiva integradora que estrutura as práticas existentes e clarifica as condições de aplicação das diferentes abordagens, contribuindo para a literatura sobre inteligência artificial e gestão estratégica e permitindo a derivação de implicações práticas para sistemas de medição que captem tanto o desempenho de curto prazo como o impacto estratégico de longo prazo.

Palavras-chave: Inteligência Artificial, Retorno sobre o Investimento, Medição do Valor da IA, Transformação Digital, Gestão Estratégica

Título: Do 0 ao 100 – Maturidade em IA e Transformação Empresarial

Autor: Ben Schütte

Table of Contents

Abbreviations

AI	-	Artificial Intelligence
ROI	-	Return on Investment
KPI	-	Key Performance Indicator
PwC	-	PricewaterhouseCoopers
TCO	-	Total Cost of Ownership
A/B	-	A/B Testing

List of Figures

1. Introduction

The thematic focus of the thesis is presented in this chapter, as it defines the rationale behind researching the issue of the return on investment (ROI) of artificial intelligence (AI) and places the topic within the existing managerial and scholarly discourse. The following sections provide the context in which AI has emerged as a highly important engine of organizational change, define the physical and theoretical issues of assessing AI-related value creation, and obtain the general research question that will inform this research.

1.1 Background and Motivation

AI has established itself as a key element of changes in companies, both in the day-to-day operations and in decision-making processes. With a growing number of investments in AI initiatives, the concern about its worth and quantifiable outcomes is growing.

The fast pace of AI across sectors has been of more significance to businesses. It is important to understand how this technology can help corporations succeed and be more efficient. The trend shows that businesses are becoming increasingly invested in AI while prioritizing efficiency and innovation (Brynjolfsson et al., 2021). In the process of measuring AI initiatives, traditional financial metrics are frequently applied to gauge the success of AI, and they fail to capture the intangible value creation, therefore, not covering all aspects (Artene et al., 2024). Against this backdrop, it is critical to find out whether firms evaluate the realistic output of AI in relation to operational and strategic outcomes.

1.2 Problem Statement and Research Gap

The determination of the ROI of AI efforts is now an important managerial challenge because organizations are investing enormous amounts of resources in such efforts as data infrastructure and talent development. Nevertheless, the quantifiable contribution of AI is, in most cases, unpredictable despite such investments. The current evaluation process is not well-designed to reflect the multidimensional and often intangible value effects produced by AI systems, making it a complicated task to make strategic choices and allocate resources (Moskalyk, 2025).

With the rise of AI and its use to influence organizational processes and capabilities, the absence of standardized and practically applicable methods of ROI measurement leaves a major gap for both researchers and practitioners. This distance is enhanced by the fact that AI technologies are rapidly developing, and no routines to assess their economic and strategic performance have been established.

1.3 Research Question and Objectives

This explores the ways organizations quantify the ROI of AI efforts in practice, and the routines, metrics, and governance forms that define this measurement. This paper aims to determine the existing practices, issues, and trends in AI-ROI evaluation by consolidating information in the field provided by specialists and precedents. As a result, the research question guiding this work is:

How do organizations evaluate the return on investment of artificial intelligence initiatives when traditional measurement approaches fall short?

1.4 Scope and Structure of the Thesis

The scope of this thesis is defined by its focus on how organizations evaluate the ROI of AI initiatives in practice. The examination will focus on massive organizations and their approaches to assessing AI-related value creation on various levels of adoption and integration. It focuses on managerial practices, measurements, and systems of governance. The research aims at a purposeful qualitative approach to describe the complexity and context-dependence.

The thesis structure indicates this coverage. Chapter 2 provides a literature review of the relevant literature on AI value creation and current methods of AI-ROI measurement, which leads to a narrowed research gap. Chapter 3 records the qualitative research design, data collection, and the analysis procedure. Chapter 4 outlines the empirical results, organized in terms of aggregated dimensions and second-order themes emerging from the results of the interview analysis. Chapter 5 presents the findings with regard to the research question, introducing and developing a framework of AI-ROI measurement, explaining its implications, limitations, and future research directions. Lastly, Chapter 6 ends the thesis by providing the summaries of the main findings and contributions.

2. Literature Review

This chapter summarizes current studies on AI value generation and how AI-related returns are assessed. With organizations investing more in AI, the issue of how to quantify the results of these projects has gained significant relevance. The literature review looks at the already instituted concept-based foundations, methods, and empirical evidence on AI-ROI measurement. It also determines the gap in existing knowledge and sets the stage for the narrowed research gap at the end of the chapter.

2.1 AI Value Creation and Challenges of ROI Measurement

The way firms develop, provide, and capture value has been transformed over a short period by AI. Other than varying technological solutions, AI is not a single technology solution. AI is a general-purpose enabler, combining data, algorithms, and human knowledge to make better decisions, automate workflows, and find new sources of innovation (Bohnsack & de Wet, 2025). AI can be used to make organizations more efficient and flexible by optimizing processes, personalizing products, and making predictions (Hanelt et al., 2020). As such, AI is not only a functional tool but also a contributor to new organizational forms.

In terms of management, AI increases the organization's responsiveness to constant learning about data, seeing patterns, and feeding the information into the strategic and operational decisions (Rahman & Mehnaz, 2024). Such feedback loops allow quicker innovation processes, better customer experience, and a higher quality of decisions, which is directly connected to competitive advantage (Sjödin et al., 2021). The further AI goes into the strategic decision-making process, the more value it adds to the bottom line associated with the intangible results of knowledge accumulation, organizational resilience, and agility, factors that are in the long term the basis of competitiveness and are hard to measure by conventional metrics (Haefner et al., 2020).

These multidimensional and evolving forms of value creation illustrate why measuring the ROI of AI initiatives presents substantial challenges. Traditional ROI models are designed for technologies with stable outputs and linear performance effects and therefore struggle to capture AIs iterative, learning-driven impact (Aldoseri et al., 2024). The diversity of AI systems and the nature of their reliance on data quality and the existence of continuous learning cycles complicate the process of setting up stable foundations and attributing the results of certain AI elements (Haefner et al., 2020). Furthermore, the AI-enabled advantages like better decision quality, increased knowledge flows, or increased ability to innovate are strategic and intangible, which results in measurement gaps when financial indicators are used (Okunola & Ahsun, 2025).

In addition to technical considerations, organizational conditions also make the ROI assessment process further difficult. The lack of data, data integration, and model transparency are common issues that hinder the credibility of the performance (Aldoseri et al., 2024; Bagehorn et al., 2025). Moreover, standardized KPIs or governance systems used to align AI performance with strategic goals are absent in most organizations, leading to fragmented or inconsistent evaluation practices (Hanelt et al., 2020; Jorzik et al., 2024). These challenges prove that assessing AIs contribution to organizational value cannot be done without a multidimensional perspective encompassing financial, operational, strategic, and learning-focused indicators.

Combined, the capability of AI to create tangible and intangible value and the fact that it relies on dynamic and context-specific learning mechanisms are a major hindrance to the conventional ROI measurement (Rajesh Sura, 2025). The knowledge of these issues gives the theoretical basis for exploring current responses to AI-ROI and evaluating how organizations strive to capture the multiple types of value generated by AI.

2.2 Existing Approaches to AI-ROI

The calculation of AI-ROI initiatives presupposes knowledge of the various methodological and conceptual models used by organizations in practice. Although conventional financial metrics may offer valuable insights into quantifiable efficiency improvements, AI-based value creation goes beyond quantifiable cost or time reductions and, more often than not, involves qualitative and ability-based results (Mannepalli, 2025). Consequently, the current methods of AI-ROI integrate operational metrics with more conceptual models that obtain a picture of both financial and non-financial impacts (PwC, 2021; BDO, 2025).

In this section, a distinction is drawn between a practical measurement approach, for example, those that aim to quantify short-term performance enhancements, and conceptual frameworks, for example, those that offer systematic views on the multidimensionality of AI value creation. The subsections below present the two sides to give a unified picture of the way AI-ROI is evaluated in research and application.

2.2.1 Practical Measurement Approach: SectionAI-ROI Framework

SectionAI can offer a viable and increasingly used method of measuring AI-ROI by offering a measurement cycle of its operations that is aimed at quantifying the short-term efficiency improvements and promoting long-term capability developments. The structured evaluation process, by SectionAI, includes three steps: (1) a pre-AI baseline, (2) a focused 90-day sprint that tests the usefulness of AI in specific workflows, and (3) a post-implementation comparison of the results to identify quantifiable gains (SectionAI, 2025).

The strategy also includes a simple ROI formula that calculates the result in improvements in performance regarding AI investments:

$$ROI = \frac{(Baseline\ Cost/Time - \ Cost\ of\ AI/Time)}{AI\ Investment}$$

This formula can be used repeatedly in the quantification of financial benefits, such as the reduction of costs or the saving of time. In addition to efficiency gains in the short-term, SectionAI focuses on the long-term strategic significance of making an organization AI-native, where employees gain the skills and knowledge necessary to carry on integrating AI into work processes day by day (SectionAI, 2025).

2.2.2 Conceptual and Industry Approaches to AI-ROI

In addition to operational measurement methods, including the SectionAI approach, various conceptual models have been created to design the organizational process of assessing the financial and non-financial payoffs of AI initiatives. These frameworks offer groupings, logic of measure, and governance principles that assist organizations not only to evaluate measurably enhanced efficiency but also far less tangible outputs like innovation, learning, or cultural change. These two are part of the growing acknowledgment that AI value cannot be measured by financial metrics but must be viewed through multidimensional perspectives of evaluation (PwC, 2021).

Common financial measures used in the AI-ROI assessment include reductions in costs, time savings, increases in productivity, and increases in revenues. These metrics allow the organizations to measure the direct effect of AI on the efficiency and performance of their operations. Nevertheless, since AI is transformative and learning-focused, non-financial measures have become significant for representing the overall organizational value of AI. They consist of better employee experience, knowledge retention, decision quality, innovativeness, and agility of the company - the aspects that are hard to quantify in financial terms but are at the core of long-term competitiveness (BDO, 2025; IBM, 2025).

One managerial framework is offered by PwC, dividing AI returns into hard and soft dimensions. Hard returns are determined by actual financial gains like reduced costs, saved time, or increased productivity, whereas soft returns relate to the qualitative advantages of employee experience, skills retention, and organizational responsiveness. PwC also separates tangible and intangible inputs (examples of which are infrastructure, licenses, personnel, data quality, expertise, and training), thereby connecting financial and non-financial aspects with particular drivers of investments (PwC, 2021).

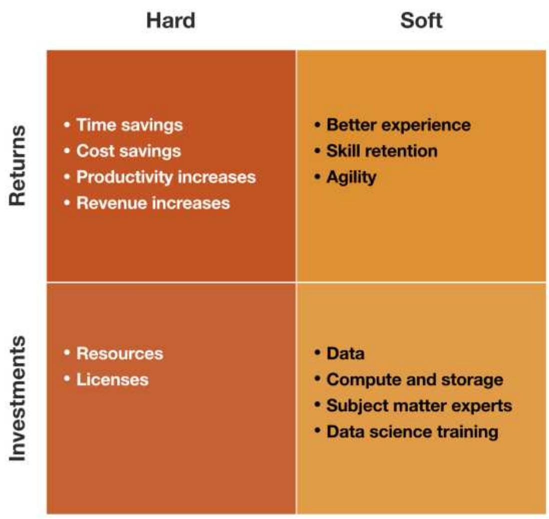

Figure 1: PwC's classification of hard and soft components of AI-ROI (PwC, 2021)

BDO extends this approach by framing AI-ROI as a multidimensional construct that cannot be captured through financial indicators alone. In BDO's view, AI also generates non-financial returns related to cultural adaptability, leadership alignment, employee adoption, and innovation readiness. These effects shape an organization's ability to integrate AI into decision-making and to build long-term capabilities that support future value creation (BDO, 2025).

Another point of view is brought to the table by IBM that focuses on governance as a requirement of quantifiable ROI. Financial returns are determined by ownership and resource distribution, but, according to IBM, non-financial benefits, including better alignment, less ambiguity, and more accountability, are essential facilitators of successful AI implementation. Companies that have developed their governing systems are more likely to achieve greater ROI as they will also incorporate both the financial and non-financial measurements into a logical approach to measurement (IBM, 2025).

Deloitte adds to these perceptions the notion of the human-agentic workforce that emphasizes the interaction between the role of humans and AI in value generation. According to Deloitte, the traditional ROI models do not work in cases where AI is integrated in autonomous or semi-autonomous systems since the joint productivity, innovation, and adaptability generated by combining human and AI cannot be quantified solely in financial terms (Deloitte, 2025). Consequently, the ROI measure is progressively becoming increasingly dependent on the incorporation of both financial metrics and qualitative ones concerning learning, flexibility, and workforce reshaping (Deloitte, 2025).

In all these methods, three common principles are seen.

1. AI-ROI measurement should be based on financial and non-financial metrics in order to represent the entire range of value creation
2. Companies need to leverage the use of time perspectives, which capture the learning nature of AI and the time required to achieve intangible gains
3. The right ROI measurement relies on governance structures in line with the alignment of financial measures and the qualitative indicators to the strategic goals

All of these principles in combination emphasize the shortcomings of existing ROI models and the requirement of holistic, context-specific assessment frameworks that acknowledge the intricacy of AI-based transformation.

2.3 Industry Practices in Measuring AI-ROI

Practices in the industry also help to gain more information concerning how organizations assess the business value of AI initiatives in actual contexts. Even though academic and consulting models describe systematic methods of measuring AI-ROI, corporate reporting illustrates how companies can communicate, measure, or estimate the value of AI investments in the field. The following examples demonstrate the methods used by large organizations of various industries.

Three large companies in various sectors, namely ***Allianz Group*** (financial sector), ***Mercedes-Benz Group*** (automotive industry), and ***IBM Corporation*** (technology industry), were chosen as representative cases. This comparison across industries enables viewing the differentiation in the approaches of organizations with different business models to evaluate the value of AI in businesses and report on their findings in annual reports.

In the ***Allianz Group*** Annual Report of 2024, AI is brought out as part of a bigger strategy of enhancing productivity, streamlining operations, and risk assessment of the organization with the help of the so-called "Business Master Platform". The company links AI investments to increased customer satisfaction and operational performance and tends to cite Net Promoter Scores (NPS) and internal productivity indicators as a proxy measure of value creation. Yet, Allianz does not disclose any quantifiable ROI measures of its AI initiatives, and financial implications are hidden in the general digitalization or technology budgets (Allianz Group, 2025).

On the same note, the ***Mercedes-Benz Group*** Annual Report of 2024 highlights AI as a source of innovation and efficiency throughout its digital operating system (*MB.OS*). AI is associated with predictive maintenance, customer personalization, and process automation, though its financial results are presented qualitatively. According to the company, digital initiatives increase efficiency and resilience, but no ROI or payback numbers are mentioned. AI is positioned as a strategic asset for the long term instead of a short-term investment that has quantifiable financial payoffs (Mercedes-Benz Group, 2025).

IBM has more sophisticated and open communication of returns regarding AI. The IBM Annual Report of 2024 specifically mentions AI and hybrid-cloud technologies as two of the pillars of its growth strategy and links them to the quantifiable business results. According to the company, its generative-AI book of business has a cumulative revenue of USD 5 billion and credits USD 3.5 billion in interior productivity savings since 2023 to AI-powered automation endeavors (IBM Corporation, 2025). IBM's reporting framework connects AI investments directly to operational indicators such as cost reduction, time savings, and improved accuracy.

Taken together, the cases show that firms vary significantly when it comes to the measurement and reporting of the value of AI programs. Although quantifying ROI is becoming less common practice in most organizations, they are increasingly using proxy metrics that bring together both financial metrics (e.g., cost savings, revenue impacts) and non-financial metrics (e.g., innovation potential, customer satisfaction). Such practices represent the prospects and constraints of the existing corporate reporting, as well as emphasize the necessity of more uniform and open methods of assessing AI-involved returns.

2.4 Organizational Perspective on AI-ROI: Three Layers of Value Creation

Continuing the previous discussion of conceptual and practical methods of measuring the AI-ROI, this section will look at how the value is created with AI in the first place. Existing studies emphasize that AI produces outputs at various levels, including organizational building blocks, operational efficiency, and strategic innovation (Raisch & Krakowski, 2021). To organize these different contributions, the following subsections differentiate between three intersecting layers of AI-enabled value creation: AI fluency as an organizational capability, internal AI use cases that generate operational improvements, and external AI use cases that facilitate new products, services, and business models. This approach offers the conceptual framework of the multidimensionality of AI-ROI.

2.4.1 ROI from AI Fluency as an Organizational Capability

The AI fluency is the overall capacity of employees and teams to comprehend, communicate, and reasonably utilize AI tools in their day-to-day activities. It is an organizational strength that can define the ability of firms to leverage AI potential into business practicalities (Disco, 2025). The more AI is incorporated into organizational processes, the less resistance is caused, the more a company can experiment with AI, and allow employees the opportunity to identify and utilize AI-assisted problem-solving opportunities (Mikalef et al., 2023).

In addition, AI fluency can improve the quality and speed of decision-making. With knowledge and comprehension of the AI systems, employees are in a better position to interpret outputs, risk assessment, and incorporate the insights into the strategic and operational processes, leading to a greater responsiveness (Deloitte, 2025).

Whereas fluency does not normally lead to immediate financial benefits, it creates the grounds on which it would be possible to record ROI benefits. Research underlines the fact that firms that possess high digital and AI competence experience greater performance impacts as a result of automation, predictive analytics, and AI systems that are customer-facing (Babina et al., 2023). In this sense, AI fluency can be understood as a long-term investment in human and organizational capital that indirectly shapes financial outcomes by enabling effective use-case execution and accelerating innovation cycles.

2.4.2 ROI from Internal Use Cases: Automation and AI Agents

Internal use cases refer to the operational level of AI value creation. These applications are towards automation, prediction, and optimization of the current processes, and they tend to create quantifiable efficiency (Davenport & Ronanki, 2018). Examples are automated document processing, AI-assisted forecasting, automation of customer service, or autonomous agents, with which regular activities are implemented. The results of such applications can be cost structure improvements, cycle time, accuracy, or productivity, which is often measured in financial terms (Patrício et al., 2024).

The conceptual frameworks, like the differentiation between hard and soft ROI of PwC and the quality of governance promoted by IBM, indicate that the internal ROI is not only presented through the technological performance but also organizational readiness (PwC, 2021; IBM, 2025). Clarity of KPIs, quality data infrastructures, and ownership frameworks can help firms trace operational gains to individual AI initiatives.

However, financial value isn't the sole operational value. Internal use cases also produce non-financial deliverables that strengthen organizational capacity, e.g., better work experience or better cross-team collaboration. These non-material impacts may have an indirect impact on the financial returns through building up adoption rates and facilitating more extensive change efforts. Consequently, internal use cases constitute the interface between lower capability development and the strategic value creation.

2.4.3 ROI from External Use Cases: AI Products and New Business Models

External use cases are AI strategic value creation in the form of new products, services, and business models. It can be personalization engines, AI-based digital platforms, predictive service offerings, or completely new solutions based on AI. Such programs go beyond in-house productivity and affect the manner in which companies compete, differentiate, and expand. The strategic value is usually based on new revenue streams, a better customer experience, gaining a better position in the market, or the opportunities to be innovative in the long term, as the Reports reveal in section 2.3 (IBM Corporation, 2025; Mercedes-Benz Group, 2025).

In contrast to operational use cases, external applications are more uncertain, have longer lifespans, and are more dependent on ecosystem dynamics (Adner et al., 2019). The adoption in the market, regulatory environment, availability of data, and integration with partners determine the success of AI-based offerings in creating value (OECD et al., 2025). This is how we see corporate reporting, where companies tend to report the results of external AI projects qualitatively, citing increased innovation, customer relationships, or personalization, since attributing it to financial metrics is not simple and is affected by a variety of interdependent variables (Valantic & Handelsblatt Research Institute, 2025).

Despite these obstacles, external applications can be the most critical strategic AI integration. They enable companies to develop new types of customer value, diversify to adjacent markets, and change the competitive dynamics. In that regard, external use cases are an embodiment of AI-enabled transformation: They are based on well-established organization capacities and on the operationalization of internal use cases, at the same time providing long-term strategic returns that determine the future direction of the firm (Akkol & Jurgens, 2022; Bughin et al., 2018).

2.5 What We Know and What We Don’t Know About Measuring AI-ROI

The current studies and industry publications offer an increasing but still disjointed insight into how the ROI of AI projects can be assessed. As it has been indicated, AI generates value at various levels within an organization, including the capacity to develop and operational efficiency as well as strategic initiatives. Such conceptual frameworks as provided by PwC, BDO, IBM, and Deloitte underline the necessity to combine both financial and non-financial measurements, take the role of the governance system into account, and regard the implications of AI on organizational learning and agility in the long term. Empirically, it is also observed that internal AI use cases are often associated with quantifiable benefits in the areas of cost reduction, time savings, or productivity, whereas external use cases are associated with innovation and the creation of new business models, which, however, are less quantifiable financially.

Although there are still huge gaps in the measurement of AI-ROI. The type of evaluation that is more applicable to technologies that produce dynamic, nonlinear, and intangible forms of value is traditional evaluation models. The iterative learning, quality of data, and changing performance of AI complicate the process of attribution and create issues in setting certain stable baselines to compare with. In addition, the advantages of many AI-related applications, including better quality of decisions, stronger cooperation, or greater organizational agility, are simply hard to measure and are not easily arranged within traditional accounting schemes. Consequently, the companies usually use proxy measures or qualitative indicators to assess the results of the AI, especially in the complex workflows that include several systems.

The discrepancies in reporting on the value of AI between firms in the industry are also noticeable. Although other companies like Allianz and Mercedes-Benz demonstrate strategic advantages and cost reduction, clear ROI values are seldom published. Even more sophisticated adopters, such as IBM, disclose only some financial metrics, which are followed by explanations in qualitative language of the overall business impact of AI. This is even though standardized measurement practices are still wanting, which shows the need to have stronger, comparative, and situation-specific practices of measuring AI returns.

Cumulatively, these points suggest that despite the useful information that current frameworks can offer, it is yet to be demonstrated that they may fulfill a sufficient basis for AI-ROI metrics in practice. The literature does not provide empirically based routines incorporating financial, operational, and strategic metrics, and industry practice is still based on conglomerate and frequently qualitative methods of assessment. The restrictions indicate the necessity of additional studies that would analyze the way organizations in reality evaluate the AI-ROI, and which metrics, processes, and governance systems contribute to the reliable assessment.

2.6 Summary and Refined Research Gap

The previous sections have indicated that AI generates organizational value by offering several interdependent levels. At the capability level, AI fluency allows organizations to apply AI tools in everyday work, improve the learning cycle, and establish the grounds that support the process of value delivery in the long term. Internal use cases of AI create a quantifiable efficiency at the operational level due to automation, prediction, and optimization of processes. The external AI case studies at the strategic level are used in support of new products and services, as well as business models that redefine competitive positioning. These layers demonstrate that AI has both financial and non-financial results, which confirms the multidimensional nature of AI-driven value creation.

The literature further notes that the current strategies of quantifying the AI-ROI are indicative of this multidimensionality. Operational models, like the SectionAI model, offer feasible protocols for setting baselines, short-term evaluation, and measuring the efficiency impacts. Several conceptual models developed by PwC, BDO, IBM, and Deloitte focus on the hybrid types of measurement logics, that is, on financial metrics and qualitative measures of learning, innovation, and governance. In spite of these developments, the existing methods are still in a state of disunity and are generally not empirically validated. They give a sense of what to measure, but do not give much insight into how firms are doing the assessments in practice.

This ambiguity is further demonstrated through insights into corporate reporting. Although more companies are talking about the strategic significance of AI, there is a paucity of explicit ROI. In most cases, companies are estimating the worth of AI initiatives based on qualitative descriptions, proxy measures, or single financial measures, which strengthens the distance between theoretical models and actual practices in measurement.

Even though the literature defines the dimensions in which AI creates value and includes a discussion of the different ways of measuring it, there is little empirical evidence regarding how organizations actually apply ROI evaluation in their practice. In particular, it is still not clear what measures, practices, and administrative frameworks companies use to evaluate AI projects, how they overcome the problem of attribution, and how financial and non-financial outcomes are incorporated into the overall assessment procedures. The gap should be bridged to optimize knowledge of the AI-ROI measurement and facilitate organizations to align measurement practices to the strategic and operational realities of implementing AI.

3. Research Design and Methodology

This chapter describes the methodology of the research. It starts with the explanation of the qualitative research method used in Section 3.1, and then details the expert sampling technique and data collection process used in Section 3.2. The analytical process of the arrangement and interpretation of the interview material is subsequently described in Section 3.3. These elements, when combined, create the methodological framework of empirical results presented in the next chapter.

3.1 Qualitative Research Approach

In the present Thesis, a qualitative research design is taken as a tool to embrace the complexity and situational character of organizational practices connected to AI-related ROI measurement. This course of action has its basis in the fact that organizational realities are social constructs that can only be comprehended based on the meanings and interpretations of actors involved. Semi-structured expert interviews act as the primary source of data so that the processes of exploration of experiences, routines, and the logic behind them may be conducted without predefined constructs.

The qualitative design is inductive as it enables analytical perceptions to be made out of the empirical material and not because of prior theoretical assumptions. The research method implies a cyclic movement between data analysis and conceptual elaboration, which will allow building a subtle comprehension of how organizations weigh AI projects. This method is deemed to be suitable for exploring complex, dynamic, and under-theorized phenomena like AI-ROI measurement.

3.2 Expert Sampling and Data Collection

The use of a purposive sampling methodology was in place to make sure that the empirical content covers a wide scope of organizational experiences, as far as the assessment of AI-based initiatives is concerned. The sampling strategy involved focusing on innovation specialists who are employed in large companies, and consultants who have experience in implementing AI were also brought in to offer more external insights. This mix was chosen as a way of incorporating external advisory thinking as well as internal practitioner thinking, and make it possible to gain a more holistic view of the practices of ROI measurement as perceived and practiced in various corporate environments.

The personal network and selected outreach via LinkedIn provided the possibility to screen potential subjects because the latter allowed reaching people who are actively interested in AI strategy, digital transformation, or the evaluation of value creation using advanced analytics. This type of outreach was useful in bringing in specialists with pertinent roles and decision-making in their respective settings. All the interviews were scheduled one by one, and they were held online over Microsoft Teams to facilitate the geographic dispersion, minimize logistical limitations, and maintain consistency in the environment of the interview across the cases.

Each interview was recorded and transcribed with the consent of the participants. The transcription process made sure that the information in the discussions remained intact and formed a solid foundation of repetitive analysis of how professionals defined routines, interpretations, and evaluation practices associated with AI-ROI measurement. Eased rules of transcription were used as much as they enhanced clarity of analysis without distorting the accounts of the experts. In cases where interviews had been conducted in German, transcripts were later translated into English.

In order to protect confidentiality and respect ethical principles of academic research, all respondents were anonymized. Professionals and their organizations are identified by distinct anonymized identifiers. The interviewed experts have been described in detail in the Table below.

Nr.	Position	Industry	Company Size (employees)	Expert ID
1	Senior Head of Digital & AI	Management Consulting	> 35,000	E1
2	Partner & Head of Technology Europe	Business Services	> 300,000	E2
3	Senior Principal for Digital Strategy	Management Consulting	> 35,000	E3
4	Business Transformation Lead	Consumer Goods	> 60,000	E4
5	Team Lead of Digital Transformation Office	Financial Services	> 25,000	E5
6	Senior Manager for Digital & AI Transformati	Energy and Utilities	> 20,000	E6
7	Chief Data & AI Officer	Infrastructure	> 2,500	E7
8	Global Head of AI-Transformation	Consumer Goods	> 5,000	E8

Figure 2: Overview of Interview Partners (own illustration)

3.3 Analytical Procedure

The analysis of the qualitative data was based on an inductive and interpretative process guided by the Gioia methodology that offers a systematic way of deriving conceptual understanding of a qualitative text (Gioia et al., 2013). The analysis was carried out in a logical series of abstraction steps following this orientation.

First, the transcripts of all interviews were read line by line to distinguish among informant-focused expressions and terms that are related to the assessment of AI initiatives. These descriptive items were recorded as the original *First-Order Concepts* with the principle of not translating the terms used by the participants in the initial stages of the analysis (Gioia et al., 2013).

Second, the newly developed *First-Order Concepts* were compared in cross interviews to determine convergences, divergences, and patterns. By this comparison and contrast, statements closely related conceptually were then grouped into more abstract Second-Order Themes. This action represents the abductive reasoning approach in the iterative model put forward by Magnani and Gioia (2023), where inductive pattern recognition is augmented with selective interaction with the extant theory.

Third, the Second-Order Themes were further generalized into *Aggregate Dimensions* that exist on a higher level of conceptual arrangements in the data. These dimensions reflect patterned logic that describes organizational values of perception and evaluation of AI-related ROI.

The analysis was also repeated over multiple steps, where empirical material was involved, newly discovered categories, and other perspectives applicable in the analysis, which guaranteed the rigor and transparency of analyzing raw data for the development of conceptual interpretation (Gioia et al., 2013). The data structure that will be the result of this will consist of the *First-Order Concepts, Second-Order Themes,* and *Aggregate Dimensions* and will be presented in the Findings chapter of the report as the analytical reasoning behind the study.

3.4 Methodological Limitations

The methodology used in this research implies a number of limitations related to qualitative, interview-based inquiry. The empirical material is based on the subjective descriptions of experts and may therefore be affected by biases linked to their job positions, professional experience, and the situational context in which the interviews were conducted. In addition, the composition of the sample restricts the range of perspectives represented and may not fully reflect the diversity of industry practices in AI return on investment evaluation. Moreover, the dynamic and rapidly evolving nature of artificial intelligence technologies and related measurement practices suggests that certain insights may change over time, thereby reducing their long-term stability. These constraints do not call into question the exploratory character of the study but instead delineate the boundaries within which the findings should be interpreted.

4. Empirical Findings

This chapter presents the empirical data from the interviews with experts. The analysis is organized in four aggregated dimensions (Sections 4.1 - 4.4) that summarize the key trends found in the data. The sub-sections in each aggregated dimension give the second-order themes in each dimension, which give a more detailed narrative of particular practices, mechanisms, and constraints associated with AI value creation and AI-ROI measurement.

4.1 Perceived AI Value Creation

The empirical study suggests that AI is not viewed as a single source of value but is more of a multidimensional mechanism through which organizations anticipate better performance outcomes within the various organizational levels. Throughout the interviews, the creation of AI-related value was continually talked about in broader scopes other than limited cost-saving aspects, but rather as a longer-term contribution to the effectiveness, competitiveness, and development of an organization.

The summary of the perceived AI value creation is presented in the form of the aggregated visual structure at this point of the analysis. *Figure 3* shows that there are several first-order concepts that are clustered into four second-order themes that compose the aggregated dimension Perceived AI Value Creation. The figure is used as an analytical anchor of the dimension and gives an overview of the value logics realized within the data of the interviews.

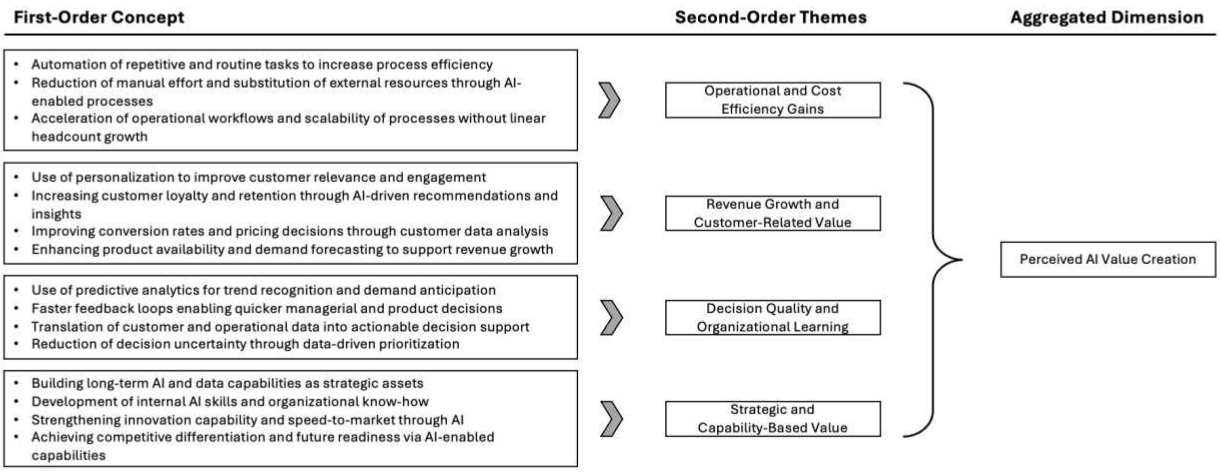

Figure 3: Aggregated Structure of Perceived AI Value Creation (own illustration)

The results imply that AI projects are connected with a variety of overlapping value logics. These value perceptions cut across operational gains, customer-related gains, advancements in decision making, and developing strategic capability. Notably, these sets of value logic were not termed as exclusive. They were instead described as complementary to each other and mutually reliant, as there was an expectation that AI can have an impact on efficiency, growth, learning, and strategic positioning.

On the operational level, AI was understood as the tool that would enhance efficiency and scalability through automating the routine tasks and increasing the speed of the current operations. Meanwhile, AI was also credited with value in facilitating revenue increase and customer-related results, especially in personalization, enhanced customer knowledge, and data-driven optimization of the product and pricing decisions. Outside these more direct impacts, AI was referred to as improving the quality of decisions and organizational learning by facilitating predictive insights, accelerating feedback loops, and making decisions under uncertainty with more insight.

Along with short- and medium-term performance impacts, there was a clear line of perceived value concerning the strategic role of AI in the development of long-term organizational capabilities. Investment in data infrastructure, AI expertise, and organizational experience was positioned as infrastructural resources that help in the ability to innovate, be future-ready, and have a long-lasting competitive advantage. Such capability-related value perceptions were often disaggregated with respect to short-term financial returns but were related to the long-term.

Combined, all these findings lead to a multidimensional perspective of AI value creation in organizations. AI is seen to enhance organizational performance in the operational, customer-facing, managerial, and strategic sectors, cutting across various time horizons. It is based on this that the four second-order themes: *operational and cost efficiency gains (4.1.1)*, *revenue growth and customer-related value (4.1.2)*, *decision quality and organizational learning* (4.1.3), and s*trategic and capability-based value (4.1.4)* were grouped to create the broad dimension of Perceived AI Value Creation.

In the coming subsections, each of these second-order themes is discussed in detail, explaining how each of the value dimensions is expressed and substantiated.

4.1.1 Operational and Cost Efficiency Gains

One of the most direct and concrete perceived AI value sources was operational efficiency, which was revealed in several of the interviews. The concept of AI was always presented as an opportunity to automate repetitive and routine processes, minimize the work of manual labor, and accelerate the speed of the process. The efficiency gains were mainly tied to the operational units and support functions, where the effective implementation of the AI-based solutions could be deployed due to the standardization of processes and large volumes of tasks.

Interviewees highlighted that AI helps to replace manual processes with computerized workflows, which leads to the speedier implementation and reduced cost of operations. One expert noted that *"[...] a lot of the value cases we see are simply about removing manual steps and replacing them with AI-supported workflows that run much faster and at lower cost"* (E1). Here, the efficiency benefits were often associated with the minimization of external resources, in such aspects as analysis, reporting, or content-related operations that had been previously supported externally.

In addition to automation, AI was also viewed as one of the key scaling levers. Some of the interviewees pointed out that AI can help organizations to produce more output without a similar increase in headcount. As one respondent explained, *"once the model is in place, you can scale the process almost indefinitely without having to add more people [...]*" (E2).

All in all, operational and cost efficiency benefits represent one of the grounded dimensions of perceived AI value. These are gains that are defined by automation, speed, and scalability of operations, and which are commonly considered the most easily accessible point of entry to AI implementation in companies.

4.1.2 Revenue Growth and Customer-Related Value

In addition to efficiency-focused value perceptions, AI was always aligned with revenue increase and customer-related performance. Throughout the interviews, value creation was often explained by the fact that AI creates more profound customer insights and converts them into more specific, data-oriented commercial activities.

Some interviewees highlighted that AI would allow a more granular perspective of customer behavior, which was viewed as a conditioning factor to personalization and loyalty-building programs. Expert 4, who is working in the Consumer Goods Industry, explained this logic as follows:

"The real value comes from understanding customers much better than before. [...] With AI, we can see patterns in behavior that we could not identify manually, and this allows us to personalize interactions in a way that actually makes a difference for loyalty and repeat purchases." (E4)

The personalization was presented as a tool to increase the rates of conversion, frequency of purchases, and relations with the customer in the long term. It was also supported by another interviewee who pointed out that AI-driven customer understanding is valuable to make better-informed commercial choices throughout the customer journey: *"If customer data is used correctly, AI helps to offer the right product at the right moment, which clearly affects revenue performance*" (E7).

Along with personalization, AI was also seen as a facilitator of revenue increase due to better demand prediction and the availability of products. The interviewees have indicated that improved predictions lessen the lost sales and allow improved planning decisions. One respondent noted that *"[...] more accurate demand forecasts help us avoid stock-outs and overproduction, which has a direct impact on sales and margins"* (E8). These were mostly said to be incremental and scalable, especially where there is customer-facing and high volume.

Combined, revenue growth and customer-related value are a specific facet of perceived AI value creation. The concept of AI was perceived as a facilitator between improved customer insight and commercially applicable outputs, which supplements the increase in operational efficiency with value additions, focused on growth.

4.1.3 Decision Quality and Organizational Learning

AI was always viewed as a source of higher-quality decisions and increased organizational learning. Throughout the interviews, value creation was not linked with the fully automated decision-making process but with the capacity of AI to expand the information base of human decisions. In this regard, AI was represented as the support mechanism that eliminates uncertainty and allows the manager to make more informed decisions.

Several interviewees highlighted that AI is value-generating because it detects patterns and signals that would not have been revealed in large and complex datasets. This was described as fundamentally changing how decisions are discussed internally, as *"[...] decisions are no longer based only on experience or intuition, but on a much richer set of data-driven insights [...]"* (E4). These effects were viewed to be of special interest in settings where the volumes of data were large and the environment was dynamic.

In addition to personal choices, AI was also viewed as the one that facilitates quicker and more organized learning at the organizational level. Interviewees emphasized that AI reduces feedback times by giving a prior indication of whether assumptions or actions were right or not. One respondent noted that *"[...] with AI, we see much earlier whether a decision works or not,* *which allows us to adjust much faster*" (E7). This approach toward learning was best expressed in connection with experimentation and testing hypotheses, within which AI was said to make cause-and-effect correlations more explicit:

"AI helps us understand why certain outcomes occur. You can test assumptions, observe the results, and then refine your decision logic based on actual data rather than assumptions." (E5)

Notably, the use of AI in decision-making was always presented as a complement to human judgments as opposed to a substitute. The ultimate accountability was defined to stay with human decision makers, especially in an uncertain or complicated situation. All in all, decision quality and organizational learning are a unique aspect of perceived AI value creation whereby AI facilitates more informed decisions and the learning process at either an individual or organizational level.

4.1.4 Strategic and Capability-Based Value

In addition to instant performance impacts, the general view of AI was of strategic and capability-based value. Throughout the interviews, AI initiatives were explained as long-term ventures that lead to the development of organizational capabilities and not only provide short-term returns. In this regard, value creation was linked with building stronger technological and organizational platforms to be used in the future to attain competitiveness.

Interviewees pointed out that the investments in data infrastructure, AI models, and skills are viewed as the preconditions of a continued value creation. In industries where competition and consumer-facing are a major concern, AI-related capabilities were explained as having the advantage of quicker innovation and organizational flexibility. One interviewee noted that *"[...] once you have the data and AI capabilities in place, you can develop new use cases much faster and react more quickly to changes"* (E4).

Such investments were emphasized as long-term and cumulative on the one hand, about organizational learning and internal know-how. As opposed to concentrating on individual applications, AI was characterized as establishing features that can be exploited in a variety of areas in the long run:

"AI is a long-term journey. You build experience with data, models, and processes, and over time this becomes a capability that the organization can use in many different areas." (E6)

Notably, these value perceptions based on capabilities were frequently reported to be not associated with short-term ROI expectations. Interviewees also admitted that these investments might not be immediately followed by quantifiable financial outcomes, but were deemed necessary in the long term, irrespective of the competitiveness and resilience.

4.2 Measurement Design and Metric System

The implementation of AI value is strongly associated with the way AI initiatives are measured and assessed in organizations. Throughout the interviews, measurement was not characterized as an activity with a purely technical focus but has been developed as a multidimensional design issue that combines analytical rigor, business cognizance, and organizational viability. Consequently, the AI-ROI measurement was defined as a process of converting the perceived value into action-based and decision-making insights.

The aggregated structure of AI measurement practices is summarized in *Figure 4* and shows the way several first-order concepts connected with experimentation, metric design, and financial evaluation fall into three second-order themes, which combine and make up the combined dimension *Measurement Design and Metric System*. The figure offers a general view of how organizations organizationally set up the evaluation of AI initiatives along the analytical and managerial levels.

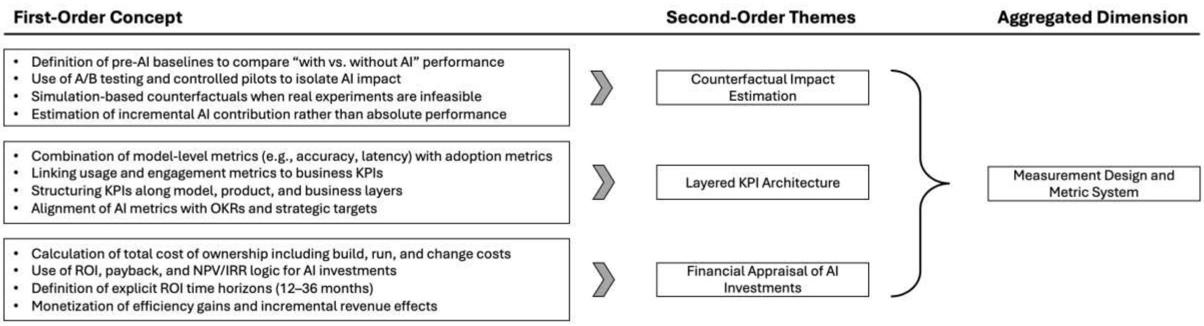

Figure 4: Aggregated Structure of Measurement Design & Metric System (own illustration)

The results indicate that AI measurement practices are defined by a high focus on identifying AI-specific influence instead of measuring the overall performance outcomes. The interviewees kept referring to the significance of counterfactual thinking with the intention of getting to know how things would have been different without AI. Simultaneously, measurement methods were explained as stratified systems that interface the technical model performance and the adoption metrics and business-level KPIs, making AI outputs related to the organizational results.

Besides operational and analytical measures, financial appraisal became one of the key elements of AI measurement. Interviewees explained that they use standard investment appraisal logics, including ROI, payback duration, and total cost of ownership (TCO), to assess

AI initiatives. Nonetheless, these financial tests were seldom implemented alone, but rather they were incorporated into comprehensive sets of metrics that capture an incremental impact, time horizon, and uncertainty.

Collectively, the findings of this paper suggest that AI measurement is not treated as an individual measure or method of evaluation but as a carefully planned system. This framework is a combination of counterfactual impact analysis, layered KPI designs, and financial evaluation tools to guide decision-making on AI investments. It is based on this that the three second-order themes: *Counterfactual impact estimation (4.2.1)*, *layered KPI architecture (4.2.2)*, and *financial appraisal of AI investments (4.2.3)* were summarized as the top-level dimension of Measurement Design and Metric System.

The subsections that follow present each of these second-order themes separately in detail, explaining how organizations implement AI measurement and how AI-related performance is translated into decision-relevant measures.

4.2.1 Counterfactual Impact Estimation

Throughout the interview responses, counterfactual impact estimation came out as one of the fundamental principles according to which the value contribution of AI initiatives should be estimated. Instead of measuring performance outcomes against the absolute results, organizations concentrated on comparing results with and without AI to confound the incremental effect that can be attributed to AI components. This counterfactual reasoning was defined as critical of credible AI-ROI evaluations.

The definition of pre-AI baselines featured strongly as one of the essential conditions of such assessments, as pointed out by interviewees. The absence of an unequivocal point of reference meant that assigning performance changes that were observed to AI was problematic. One interviewee noted that *“[...] without a clear baseline, it is almost impossible to say whether AI actually created value or whether something else changed at the same time”* (E2).

Controlled experiments, especially A/B testing and pilot designs, were very often mentioned as a choice of the counterfactual estimation. These strategies were said to allow for a stricter isolation of AI effects, particularly in digital and customer-facing settings. As one interviewee stated, *“A/B testing gives us the confidence to say that an improvement really comes from the AI and not from other changes”* (E4).

Nevertheless, the interviewees have also admitted that such experimental designs are not always possible. Counterfactuals were explained as approximated by simulation or modeled situations in operational settings when it was not possible to generate parallel control groups or where the processes were of low frequency and highly integrated. In such circumstances, the AI effect was measured by the difference between observed results (with AI) and hypothetical model-based estimates of the performance of the process in the absence of AI. This method was termed to be less accurate than controlled experiments, but it had to be adopted in real-world conditions.

Overall, the concept of counterfactual impact estimation was considered a prerequisite but situation-specific AI value measurement basis, which demands a methodological sacrifice between analytical rigor and feasibility.

4.2.2 Layered KPI Architecture

Throughout the interviews, the measurement of AI was frequently defined as the need to have a layered KPI structure that links technical performance indices with metrics of usage and business results. Instead of focusing on one indicator of success, organizations highlighted the importance of integrating several layers of metrics to convert the performance of AI into decision-specific information.

On the bottom tier, interviewees mentioned model-based measures, including accuracy, precision, or latency. These indicators were deemed to be needed to measure technical functionality, but not the ones to measure organizational value. As one interviewee explained, *"model metrics tell you whether the system works technically, but they do not tell you whether it creates value for the business[...]"* (E2). Technical performance was therefore seen as a condition and not a measure of value in itself.

A second layer was dedicated to usage and adoption indicators that were reported to be essential in closing the gap between technical performance and business impact. Interviewees highlighted the importance of the creation of value by AI systems being utilized and integrated into everyday workflows. One of the respondents summarized as follows:

"You can have a very good model, but if no one uses it in practice, there will be no impact. Adoption is the key link between the model and the business outcome." (E5)

The topmost level of responses from interviewees emphasized the need to connect AI-related measures with business KPIs and business objectives, including cost savings, improved productivity, or impact on revenue. The linkage of AI-specific indicators with conventional business performance indicators was presented as a critical method of providing managerial relevance and the ability to compare the initiatives. In general, the layered KPI architectures were considered as one of the key processes of aligning AI performance into organizational value.

4.2.3 Financial Appraisal of AI Investments

Financial appraisal was defined as an obligatory but complicated aspect of AI-ROI analysis. The established logics of financial assessment, such as TCO, ROI, and payback considerations, were identified by interviewees as significant points of reference used to frame investment decisions. Simultaneously, these strategies were represented as the ones that would need to be adjusted because AI-related value tends to be realized gradually and in uncertain circumstances.

There was a high focus on the adoption of a holistic cost approach. According to the interviewees, financial assessment usually begins with analyzing the total costs of AI projects, such as development costs, operational costs, and costs of continued changes. As one interviewee explained, *"[...] you have to look at the full cost picture, not just the initial build, otherwise the numbers will be misleading"* (E3). Simultaneously, converting AI-related benefits into financial values was characterized as a challenging task, especially when savings in time or indirect performance gains in the form of benefits are involved. One industry expert noted that *"[...] saving time does not automatically mean saving money, so translating those effects into financial value is not always straightforward [...]"* (E6).

In general, the financial appraisal was considered to be an essential part of disciplining AI investment decisions, but could not alone reasonably represent the AI value. Financial metrics were characterized as giving direction, but not accuracy, and managers need to use their judgment to make sense out of AI-related returns.

4.3 Governance and Standardization Mechanisms

The interviews also indicated that the identification and appraisal of AI value are heavily reliant on the manner in which AI initiatives are controlled and standardized in organizations. Governance was not represented as a formal control role but instead constructed to be a collection of organizational mechanisms that organize the process of decision-making, align the stakeholders, and allow comparability among AI projects. Governance and standardization in this regard were seen as facilitating factors for the effective deployment of AI and reasonable ROI measurement.

Figure 5 depicts the way in which several first-order concepts were related to steering, coordination, and standardization, and how they are aggregated into two second-order themes and, in turn, make the aggregated dimension Governance and Standardization Mechanisms. The figure gives a general view of how the AI initiatives are prioritized, monitored, and aligned over the organization's arrangements.

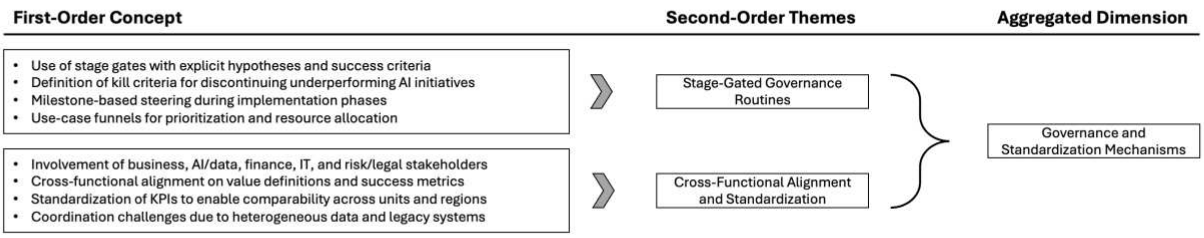

Figure 5: Aggregated Structure of Governance and Standardization Mechanisms (own illustration)

Two interrelated objectives are implied by the findings of governance practices. To begin with, organizations aim at controlling uncertainty and allocation of resources using routable decision processes. Second, they will develop common standards that provide AI initiatives with an opportunity to be compared and assessed across organizational units. These were achieved by a mix of stage-gating governance practices and cross-functional realignment and standardization practices.

Collectively, the interviews suggest that governance and standardization cannot only be related to compliance and oversight. They instead serve as integrative processes that can be used to convert strategic intentions to coordinated action and offer some framework by which AI initiatives should be evaluated. It is based upon this that the two second-order themes, *stage-gated governance routines (4.3.1)* and *cross-functional alignment and standardization (4.3.2)*, were combined into the larger dimension of Governance and Standardization Mechanisms.

4.3.1 Stage-Gated Governance Routines

Stage-gated governance routines have been explained as one of the mechanisms of driving AI initiatives in the face of uncertainty. In most interviews, projects involving AI were not normally considered using a single upfront business case. Rather, interviewees focused on iterative forms of governance where initiatives advance through a set of distinct phases with clear hypotheses and criteria of success. It was seen as a way of experimenting and, at the same time, keeping managerial control.

As noted by several interviewees, stage gates enable organizations to gauge AI projects on a case-by-case basis as well as to distribute resources on a case-by-case basis. Projects are re-examined at pre-determined points instead of making a major commitment. One interviewee explained that they *"[...] structure AI initiatives in phases, where each phase has clear* *assumptions and criteria that need to be fulfilled before moving on"* (E8). These practices were said to be especially applicable based on the unpredictable performance and changing demands of AI systems.

Moreover, the stage-gated routines were also associated with clear termination procedures besides directing the progression. Interviewees highlighted the need to establish the kill criteria to terminate non-performing initiatives and the sunk-cost effect. As one respondent summarized, *"[...] you need clear points where you honestly decide whether to continue or stop an AI use case"* (E1). In general, stage-gated governance practices were considered an effective approach to trade off risk management and experimentation in AI portfolios.

4.3.2 Cross-Functional Alignment and Standardization

Cross-functional alignment and standardization were outlined as key governance tools in terms of consistency and comparability of AI initiatives. Interviewees pointed out that AI projects are often comprised of a variety of organizational functions that include business units, data and AI teams, IT, finance, and risk / legal departments in some instances. Coherence in decision-making and plausible evaluation of ROI were seen as the benefits of aligning these stakeholders.

Some of the interviewees pointed out that universal definitions of value and success measures are the key drivers in this alignment process. In the absence of shared norms, AI initiatives were characterized as hard to compare or prioritize between units in an organization. One interviewee noted that *"[...] if every team uses different metrics and definitions, you cannot really compare AI use cases or decide where to invest next"* (E7).

At the same time, interviewees admitted that the implementation of cross-functional alignment is not easy, especially in those organizations where the data landscape is heterogeneous and legacy systems are used. The lack of coordination between units in terms of data availability, tooling, and maturity was outlined as a recurring barrier to coordination. One respondent summarized, *"bringing all functions onto the same page takes time, especially when data and systems are very different across the organization"* (E2). Overall, cross-functional alignment and standardization were viewed as critical governance mechanisms that support consistent evaluation and scaling of AI initiatives across the organization.

4.4 Barriers and Constraints to AI-Value Realization

Although the above sections have demonstrated how organizations view, quantify, and manage the value of AI-related activities, the interviews indicate that there are significant shortcomings in the transformation of such activities into actual outcomes. The interviewees highlighted several times that even in the case when clear expectations of value and designed measurement and governance systems are implemented, the realization of AI values is often hindered. These barriers were claimed to be long-term constraints and not short-term implementation problems.

Figure 6 sums up all the primary obstacles found during the interviews. The figure consolidates the several first-order challenges into three second-order themes, and combined, they constitute the dimension Barriers and Constraints to AI Value Realization. These themes, as opposed to being individual issues, reflect recurring trends that define the way the AI efforts are rolled out in reality.

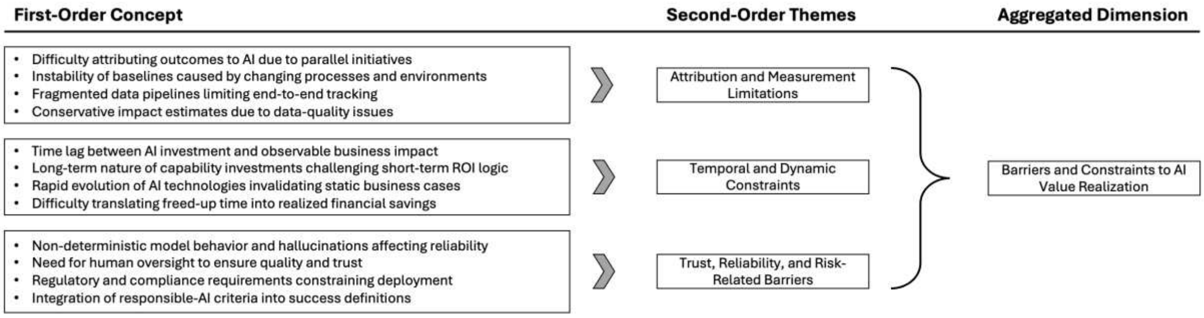

Figure 6: Aggregated Structure of Barriers and Constraints to AI-Value Realization (own illustration)

There are three broad categories of constraints: First, the *attribution and measurement problems (4.4.1)* limit the causal relationship between the observed outcomes and AI initiatives. Second, the mismatch between the short-term logics of evaluation and the long-term, changing nature of most AI investments in the form of *temporal and dynamic constraints (4.4.2)*. Third, AI use and scaling are restricted by *trust, reliability, and risk-related barriers (4.4.3)*.

In general, these results suggest that the realization of AI values cannot be explained only by the definition of value, metrics, or governance structures. Rather, it is influenced by structural and contextual boundaries that influence the perception, belief, and response of AI results in organizations.

4.4.1 Attribution and Measurement Limitations

One of the main obstacles to the AI value creation that emerged through the interviews concerns the challenges of linking the observed consequences to AI efforts. Interviewees emphasized that AI is frequently implemented in concert with other changes, like redesigning processes, introducing new digital resources, and it is difficult to causally factor these changes into it. In that case, the ability to separate the contribution of AI was termed challenging, even in cases in which measurement systems exist.

Moreover, the attribution issues were also intimately connected to turbulent or changing baselines. According to interviewees, the nature of business and underlying processes often changes AI implementation, which invalidates the pre- and post-implementation performance comparisons. This problem was especially relevant in dynamic contexts that are customer-facing. One Interviewee outlined the problem:

"It is very hard to say what exactly the AI contributed, because at the same time we change processes, interfaces, and sometimes even the business logic. The baseline keeps moving." (E7)

In addition to baseline instability, fractured data pipelines were characterized as an additional constraint on end-to-end monitoring of AI value. Interviewees mentioned that the problem of data quality and the lack of connection between systems tends to lead to limited impact estimation. Consequently, the attribution was conceptualized as a structural constraint instead of a technical issue that had to be resolved to limit the accuracy and reliability of the AI-ROI.

4.4.2 Temporal and Dynamic Constraints

Time and change constraints were found to be another obstacle to AI value realization. Interviewees repeatedly pointed to the fact that AI investments usually take a long time before one can see a business impact. This delay was termed as causing tension with quarterly assessment cycles and ROI requirements, especially in the organizations that were used to quick performance assessment.

Some interviewees emphasized that there are numerous long-term AI projects, particularly those that are aimed at the creation of data infrastructure or organization-level capabilities. Value, in these situations, was said to develop over time and not through discrete results. One of the interviewees put this challenge in a nutshell by saying:

"With many AI initiatives, you invest upfront and only see the effects much later. This makes it difficult to evaluate them using short-term ROI logic." (E3)

Besides the delays in time, interviewees also cited the fast development of AI technologies as another limitation. Business cases and assumptions were reported to be becoming obsolete at a rapid rate as the models, tools, and capabilities change. Consequently, time and dynamic limitations were presented as restricting the permanence and trustworthiness of AI values evaluations across time.

4.4.3 Trust, Reliability and Risk-Related Barriers

The issues of trust and risk were also outlined as a significant limitation on the realization of AI value. Interviewees pointed out that the behavior of a system in non-deterministic ways that include inconsistent output constrains the degree to which AI systems can be trusted in operational and decision-making situations. This consequently made the usage of AI conditional upon perceived reliability, but not necessarily upon technical performance.

Some interviewees emphasized that human control was still needed to achieve quality and reduce the risk. This was termed to be especially applicable in controlled or high-stakes settings, where mistakes can be critical. One interviewee noted that *"[...] as long as the system can produce unpredictable results, you need humans in the loop to monitor and correct outputs"* (E5).

Besides, compliance and regulatory requirements were reported to reduce the amount of AI, in terms of the scope of deployment and value realization. One of the respondents summed this limitation in a single sentence:

"You cannot evaluate AI only based on performance or efficiency. Trust, compliance, and risk considerations fundamentally limit how much value you can actually realize." (E2)

Altogether, the notions of trust, reliability, and risk-related hindrances were described as structural boundaries that shackle the value of AI realization, despite high technical performance and usage potential.

5. Discussion

In this chapter, the discussion will provide the results of the research question and place them in the wider context of the debate on the AI-ROI measurement. The synthesized answer to the research question is introduced in section 5.1, as it presents a proposed framework as a guideline for AI-ROI measurement. Based on this, section 5.2 elaborates on the interpretation of the framework, exploring the four measurement quadrants and the practices that accompany them. Finally, in Section 5.3, the limitations of the study are described, and the future directions of research are mentioned.

5.1 Answering the Research Question: A Framework for AI-ROI Measurement

The research question of the thesis aimed for the measurement of the ROI of AI initiatives in organizations, and the measurement types that could be appropriate in various phases of the AI lifecycle. The results of the empirical investigation suggest that AI-ROI measurement cannot be effectively condensed to one universal metric and approach, which has been highlighted in prior research (Haefner et al., 2020). Rather, it can be inferred that measurement practices are dependent on the extent of AI implementation and the type of anticipated value outputs.

In accordance with the synthesis of expert interviews, supplemented by literature and the examples of annual reports, a framework for measuring AI-ROI is presented as a systematic approach to the research question (*Figure 7)*. The framework does not purport to offer a clear answer to the issues of AI-ROI measurement. Instead, it provides a coherent typology according to the current measurement practices.

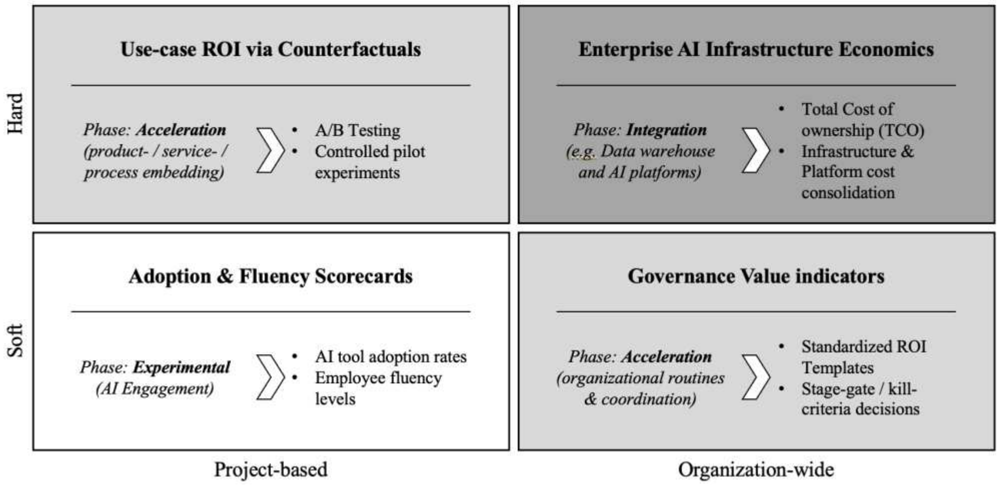

Figure 7: Practice-Based Framework for AI-ROI Measurement (own illustration)

The model has been designed in two dimensions. The first dimension distinguishes between project-based AI initiatives and organization-wide AI deployments, as it represents the level of value measurement on an individual use-case basis or at the enterprise scale. The second dimension identifies soft and hard ROI, which identify the disparity between qualitative outcomes and quantitative, financially oriented results. The combination of these dimensions provides four types of practices for measuring AI-ROI.

Measurement practices at the project level with soft ROI orientation are centered around adoption and fluency scorecards in which early value is realized by using, engaging, and developing abilities as opposed to individual financial payoffs. At the project level, where ROI is hard, use-case ROI through counterfactuals takes the center stage, which builds on baselines, pilots, or controlled comparisons to credit the improvements in performance to AI. As AI implementation becomes widespread within the organization and the organization is oriented towards a soft ROI, the governance value indicators become focal points, and coordination, standardization, and quality of decisions become important sources of indirect but consequential values. Lastly, at the organization-wide level and hard ROI orientation, enterprise AI infrastructure economics defines an ROI measurement system based on cost-focused logics, including TCO, and not on independent use-case returns.

In organizing AI-ROI measurement practices according to these four types, the framework is a direct reaction to the ambiguity that is demonstrated in the literature about how and where AI value can be measured (Deloitte, 2025; PwC, 2021). Instead of settling this ambiguity by using a single measure, the framework makes it clear that various measurement logics co-exist and are used for a variety of purposes. In this regard, the framework responds to a number of the open questions that were posed in the previous discussion on what is known and unknown regarding AI-ROI measurement (*section 2.5*). Nevertheless, it also conclusively states that measurement of AI-ROI is always context dependent and that there are always trade-offs between rigor, feasibility, and relevance.

The framework is therefore best construed as an integrative prism and, therefore, one that links the empirical practices to the existing theoretical debates, as opposed to an exhaustive or prescriptive model. The main contribution it makes is making the diversity of AI-ROI measurement practices visible and giving it a systematic ground on which to interpret how organizations manage the tension between experimentation, accountability, and strategic integration in the context of measuring the value of AI.

5.2 Implications for AI-ROI Measurement Practice

Section 5.2 explains why the suggested framework is significant to the practice of AI-ROI measurement by considering the four approaches to measurement one by one. The major basis of the discussion is the interview findings, and is supported by some selected information in the literature and annual reports.

5.2.1 Adoption & Fluency Scorecards

In all interviews, the initial AI programs were related to the measurement practices focused on adoption, usage, and employee fluency as opposed to direct financial results. As AIs are usually incorporated into individual tasks and workflows in experimental, project-oriented scenarios, it is too early to have reliable financial attribution. Interviewees emphasized the fact that early measurement is concerned with the fact that AI is actually used and comprehended, since any downstream value creation is perceived to be contingent upon sustained adoption (E5; E6; E7; E8).

Adoption and fluency scorecards were said to be pragmatic tools to measure such early value indicators. The most common ones are frequency of use, extent of adoption by teams, and perceived increases in task efficiency or quality. These indicators can be seen as value proxies, which means that in the absence of adequate levels of fluency and interaction, AI projects will likely remain in the stage of experimentation. This logic is replicated in annual reports and previous literature, which focus more on ensuring and enabling building, and mostly avoid explicit ROI assertions during the initial stages (Allianz Group, 2025; Deloitte, 2025; Mercedes-Benz Group, 2025; Mikalef et al., 2023).

In a practical sense, the results imply that organizations need to intentionally keep down the desire for ROI measures at experimental levels. Adoption and fluency scorecards provide a “good-enough method” of determining progress without compelling premature monetization. Organizational practices of measurement can be aligned to the realities of early AI experimentation by situating fluency as a capital investment and adoption as a condition to scale.

5.2.2 Use-case ROI via Counterfactuals

With AI projects leaving the speculative stage of experimentation and entering the operational stage of scaling, the interviews with experts also indicate a noticeable shift in measurement activities towards economically-based evaluation at the use-case level. Interviewees emphasized that when AI solutions start impacting the core processes or require a permanent investment, they are under increased pressure to provide economic justification for these efforts (E1; E3; E5; E8). During this stage, measurement moves away from the exploratory indicators and moves towards strategies that attempt to give credit to observable performance gains that can be directly attributed to the use of AI.

Throughout the interviews, counterfactual logics proved to be the most prevalent way of estimating use-case ROI (E2; E5; E6; E7; E8). Approaches like A/B testing, controlled pilots, and pre/post comparisons with explicit baselines were repeatedly pointed out as useful means of estimating the outcomes of the type of with-AI versus without-AI. Such practices have been appreciated since they allow making a more plausible attribution of the effects, including time savings, cost reduction, or quality improvements to AI use cases.

This focus on counterfactual measurement is widely in line with the results of the literature, which conceptualizes causal attribution as a condition of performing meaningful ROI calculations, as well as disclosures of annual reports, which tend to report efficiency or productivity increases without specifying the measurement design (Aldoseri et al., 2024; Allianz Group, 2025; IBM Corporation, 2025; Rajesh Sura, 2025).

In practical terms, the results indicate that counterfactual methods are the soundest method of measurement during the acceleration stage, although their use should not be comprehensive. It is suggested that organizations should only reserve rigorous counterfactual measurement for the high-impact or scalable use cases and use more straightforward baseline comparisons where methodological constraints preclude feasibility.

5.2.3 Governance Value Indicators

During the interviews, the measurement practices associated with governance were characterized by the fact that they grow more important when AI initiatives go beyond their single projects and start to cover several teams or business units. Interviewees said that at this point, value generation is not about the performance of particular use cases, but rather about how an organization can coordinate, prioritize, and direct a proliferating AI portfolio (E2; E4; E5; E7). This has led to measurement concentrating on governance value indicators that are associated with decision quality, comparability, and organizational alignment, and not direct financial outcomes.

The presence of and frequent application of stage-gate processes, clearly defined success and termination criteria, and common overarching evaluation logics are commonly cited indicators of AI initiatives. The mechanisms were positioned as facilitating the prerequisites of credible ROI measurement as they organize the process of the evaluation of initiatives, comparing them, and either magnifying or eliminating them. This logic is reflected indirectly in annual reports that focus on governance models, accountable AI models, or centralized coordination models, and do not mention ROI numbers much at this tier (Allianz Group, 2025; IBM Corporation, 2025; Mercedes-Benz Group, 2025). Similar in literature, governance and standardization are seen as paramount facilitators in coping with AI-led change, especially in the context of diffuse and indirect value (Hanelt et al., 2020; Raisch & Krakowski, 2021).

The results imply that the indicators of governance value can be used to complement the use-case-level ROI indicators. Instead of direct measures of value creation, they expand the ability of the organization to effectively allocate resources, as well as to sustain investment in initiatives that perform poorly. In this regard, governance indicators serve as a soft ROI to enable scalability and readiness to measure, establishing the circumstances within which the financial evaluation can be more rigorous during its implementation in later phases.

5.2.4 Enterprise AI Infrastructure Economics

On the organization-wide AI implementation level, the expert interviews provide an additional change in the ROI measurement to the enterprise-oriented cost logics. When AI is adopted into common platforms, data structures, or other fundamental technological platforms, interviewees highlighted that value is meaninglessly measured at the scale of individual use cases (E3; E8). Rather, there are measurement practices that concentrate on the economic consequences of having and running AI infrastructure as a basic organizational resource.

The concept of total cost of ownership (TCO) became the standard point of reference when it comes to evaluating AI infrastructure in an enterprise. Interviewees emphasized that initial development or acquisition costs should not be the only ones considered, but also current expenses in relation to running, maintenance, compliance, and continuous improvement (E4; E5; E6; E7; E8). This cost-centric view was said to be especially applicable during integration stages, which require AI capabilities to be integrated into various functions and can be shared as an investment instead of a specific investment. This reasoning is evident in annual reports, which usually construct AI infrastructure as part of more general digital or technology investment stories, but offer minimal detail on disaggregated ROI numbers (Allianz Group, 2025; Mercedes-Benz Group, 2025). The literature also conceptualizes infrastructure as a physical contribution to the value creation of AI, where cost transparency frequently goes ahead of quantifiable revenue or efficiency benefits (Hanelt et al., 2020; PwC, 2021).

In a pragmatic sense, the results indicate that enterprise AI infrastructure economics must have a different measurement logic than that of project-level ROI evaluation. Instead of associating the returns with particular applications, organizations are advised to weigh the infrastructure investments on multifaceted cost structures to facilitate budgeting, priorities, and long-term forecasting. Taking AI infrastructure as a business-wide economic commitment, organizations are able to adopt the ROI measurement that matches the integrated and long-term types of AI-enabled business operations.

5.3 Limitations and Future Research

This research has a number of weaknesses. To start with, the empirical evidence relies on eight interviews with experts, which portrays the accessibility of informed respondents on AI-ROI measurement, which is limited. A large number of potential participants were not able to provide information on detailed insights because of the issues of the infantile nature of measurement practices, reflecting the novelty of the topic and the existence of a significant gap in research. This, however, restricts the generalizability of the findings as well.

Second, the suggested framework is not empirically tested in an organizational practice. Instead of being a prescriptive model, it evolved through the synthesis of various methods of measurement found in interviews and was backed up by the literature and annual reports. The framework can therefore be viewed as an organizing lens and not a tested solution.

Future studies should then aim at testing and refining the framework in an empirical manner, say, conducting longitudinal and comparative case studies. Since the AI technologies are rapidly changing, additional research should also consider how the practice of AI-ROI measurements evolves, as the method of measuring it will have to change in line with the change in technology and organization.

6. Conclusion

This thesis was aimed at exploring the way organizations value the ROI of AI projects and how it is determined by practices and issues. The results prove that AI-ROI measurement can not be narrowed down to one metric/evaluation logic. Rather, AI creates value at various organizational levels and financial results, together with ability building, learning, and strategic placement. Consequently, companies use heterogeneous methods of measurement that differ based on the range of AI implementation and the type of anticipated value.

A framework of AI-ROI measurement was created through the synthesis of the findings on the topics of expert interviews, literature, and practices of annual reporting. The framework distinguishes present AI initiatives based on project and organization-wide, as well as soft and hard ROI orientation, meaning four possible classifications of measurements. They are adoption scorecards, fluency scorecards, use-case ROI through counterfactuals, governance value indicators, and enterprise AI infrastructure economics. Instead of having a prescriptive answer, the framework offers an integrative lens that puts the current practices in order and explains when different measurement logics are used in practice.

The research is significant to the field of study because it empirically bases the discussion on the measurement of AI-ROI and links fragmented conceptual frameworks with organizational routines. To practice, the results confirm that successful AI-ROI measurement necessitates alignment of evaluation approaches with the maturity and scope of AI initiatives, and a trade-off between rigor, feasibility, and relevance. It is indicated that measurement practices are developing together with the introduction of AI, whereby the former ones are less structured, financial, and governance-based logics.

All in all, the thesis makes it clear that AI-ROI is a changing organizational, but not a calculational, practice. With AI technologies and organizational capacity constantly changing, measurement methods have to be flexible, situational, and embedded in overall decision-making and governance.

7. Bibliography

Adner, R., Puranam, P., & Zhu, F. (2019). What Is Different About Digital Strategy? From Quantitative to Qualitative Change. *Strategy Science*, *4*(4), 253–261. https://doi.org/10.1287/stsc.2019.0099

Akkol, Ö. B., & Jurgens, J. (2022). *Unlocking Value from Artificial Intelligence in Manufacturing*. World Economic Forum. https://www.weforum.org/publications/unlocking-value-from-artificial-intelligence-in-manufacturing/

Aldoseri, A., Al-Khalifa, K. N., & Hamouda, A. M. (2024). Methodological Approach to Assessing the Current State of Organizations for AI-Based Digital Transformation. *Applied System Innovation*, *7*(1), 14. https://doi.org/10.3390/asi7010014

Allianz Group. (2025). *Annual Report | Allianz*. Allianz.Com. https://www.allianz.com/en/investor_relations/results-reports/annual-reports.html

Artene, A. E., Domil, A. E., & Ivascu, L. (2024). Unlocking Business Value: Integrating AI-Driven Decision-Making in Financial Reporting Systems. *Electronics (2079-9292)*, *13*(15), 3069. https://doi.org/10.3390/electronics13153069

Babina, T., Fedyk, A., He, A., & Hodson, J. (2023). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, *151*, 103745. https://doi.org/10.1016/j.jfineco.2023.103745

Bagehorn, F., Brimijoin, K., Daly, E. M., He, J., Hind, M., Garces-Erice, L., Giblin, C., Giurgiu, I., Martino, J., Nair, R., Piorkowski, D., Rawat, A., Richards, J., Rooney, S., Salwala, D., Tirupathi, S., Urbanetz, P., Varshney, K. R., Vejsbjerg, I., & Wolf-Bauwens, M. L. (2025). *AI Risk Atlas: Taxonomy and Tooling for Navigating AI Risks and Resources* (No. arXiv:2503.05780). arXiv. https://doi.org/10.48550/arXiv.2503.05780

BDO. (2025). *AI ROI: A pragmatic measurement guide*. https://www.bdo.be/en-gb/insights/articles/2025/ai-roi-a-pragmatic-measurement-guide

Bohnsack, R., & de Wet, M. (2025). *AI is the Strategy: From Agentic AI to Autonomous Business Models onto Strategy in the Age of AI* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2506.17339

Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The Productivity J-Curve: How Intangibles Complement General Purpose Technologies. *American Economic Journal: Macroeconomics*, *13*(1), 333–372. https://doi.org/10.1257/mac.20180386

Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018). *Modeling the global economic impact of AI | McKinsey*. https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy?utm_source=chatgpt.com

Davenport, T. H., & Ronanki, R. (2018). *Artificial Intelligence for the Real World*.

Deloitte. (2025). *Work, Reworked: Leading with Agentic AI*. https://www.deloitte.com/global/en/services/consulting/perspectives/human-agentic-workforce.html

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, *16*(1), 15–31. https://doi.org/10.1177/1094428112452151

Haefner, N., Wincent, J., Parida, V., & Gassmann, O. (2020). Artificial intelligence and innovation management: A review, framework, and research agenda☆. *Technological Forecasting and Social Change*, *162*, 120392. https://doi.org/10.1016/j.techfore.2020.120392

Hanelt, A., Bohnsack, R., Marz, D., & Antunes Marante, C. (2020). A Systematic Review of the Literature on Digital Transformation: Insights and Implications for Strategy and

Organizational Change. *Journal of Management Studies*, *58*(5), 1159–1197. https://doi.org/10.1111/joms.12639

IBM. (2025). *How to maximize ROI on AI in 2025*. https://www.ibm.com/think/insights/ai-roi

IBM Corporation. (2025). *Annual Report—Investor Relations | IBM*. https://www.ibm.com/investor/services/annual-report

Jorzik, P., Klein, S. P., Kanbach, D. K., & Kraus, S. (2024). AI-driven business model innovation: A systematic review and research agenda. *Journal of Business Research*, *182*, 114764. https://doi.org/10.1016/j.jbusres.2024.114764

Magnani, G., & Gioia, D. (2023). Using the Gioia Methodology in international business and entrepreneurship research. *International Business Review*, *32*(2), 102097. https://doi.org/10.1016/j.ibusrev.2022.102097

Mannepalli, P. (2025). Generative AI ROI Benchmarks & Performance Metrics for 2026. *Kanerika*. https://kanerika.com/blogs/roi-of-generative-ai/

Mercedes-Benz Group. (2025). *Full Year Results 2024 Mercedes-Benz Group*. Mercedes-Benz Group. https://group.mercedes-benz.com/investors/reports-news/annual-reports/2024/

Mikalef, P., Lemmer, K., Schaefer, C., Ylinen, M., Fjørtoft, S. O., Torvatn, H. Y., Gupta, M., & Niehaves, B. (2023). Examining how AI capabilities can foster organizational performance in public organizations. *Government Information Quarterly*, *40*(2), 101797. https://doi.org/10.1016/j.giq.2022.101797

Moskalyk, L. (2025). Methodological Challenges AND Conceptual Approaches to Measuring the Impact OF Artificial Intelligence on Roi. *Grail of Science*, *55*, 230–241. https://doi.org/10.36074/grail-of-science.22.08.2025.024

OECD, Boston Consulting Group, & INSEAD. (2025). *The Adoption of Artificial Intelligence in Firms: New Evidence for Policymaking*. OECD Publishing. https://doi.org/10.1787/f9ef33c3-en

Okunola, A., & Ahsun, A. (2025). *Measuring the ROI of AI-Driven Workforce Transformation Initiatives* (SSRN Scholarly Paper No. 5225996). Social Science Research Network. https://doi.org/10.2139/ssrn.5225996

Patrício, L., Varela, L., Silveira, Z., Patrício, L., Varela, L., & Silveira, Z. (2024). Integration of Artificial Intelligence and Robotic Process Automation: Literature Review and Proposal for a Sustainable Model. *Applied Sciences*, *14*(21). https://doi.org/10.3390/app14219648

PwC. (2021). *Solving AI's ROI problem. It's not that easy.* PwC. https://www.pwc.com/us/en/tech-effect/ai-analytics/artificial-intelligence-roi.html

Rahman, P., & Mehnaz, S. (2024). International Journal for Multidisciplinary Research (IJFMR). *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5054029

Raisch, S., & Krakowski, S. (2021). Artificial Intelligence and Management: The Automation–Augmentation Paradox. *Academy of Management Review*, *46*(1), 192–210. https://doi.org/10.5465/amr.2018.0072

Rajesh Sura. (2025). Measuring ROI of data and analytics programs: A framework for enterprise impact. *World Journal of Advanced Engineering Technology and Sciences*, *16*(3), 265–276. https://doi.org/10.30574/wjaets.2025.16.3.1343

SectionAI. (2025). *The Enterprise Guide to Tracking AI ROI*. https://www.sectionai.com/resource/the-enterprise-guide-to-tracking-ai-roi

Sjödin, D., Parida, V., Palmié, M., & Wincent, J. (2021). How AI capabilities enable business model innovation: Scaling AI through co-evolutionary processes and feedback loops. *Journal of Business Research*, *134*, 574–587. https://doi.org/10.1016/j.jbusres.2021.05.009

Valantic, & Handelsblatt Research Institute. (2025). *Study Digital 2030: The Rise of Applied AI*. https://www.valantic.com/en/research/digital-2030-trend-report/

What is AI Fluency? A Guide for 2026 Trends and Skills. (2025). Disco. https://www.disco.co/blog/what-is-ai-fluency-a-guide-for-2026-trends-and-skills

Appendix

Expert Interview Guide

Context: *Understand the expert's role and the organizational setting*

Q1.1: Could you briefly describe your role and how it relates to AI within your organization?

Q1.2: How long has your company been working with AI initiatives?

Q1.3: If you think about your organization's AI adoption, would you place it more in the (1) **experimentation**, (2) **acceleration**, or (3) **integration** stage?

Q1.4: What are the main goals your organization is pursuing with AI: (1) efficiency, (2) innovation, (3) customer experience, (4) sustainability, or (5) something else?

Perception of Value: *Understand how AI-generated value is interpreted.*

Q2.1: When thinking about your organization's AI initiatives, how do you personally define ***success***?

Q2.2: What types of value do AI initiatives generate in your view (financial, operational, strategic, or intangible)?

Q2.3: How do stakeholders (e.g., management, finance, or business units) differ in how they interpret "AI value"?

Measurement and ROI Practices: *Explore how organizations measure and report AI performance.*

Q3.1: How does your organization typically measure whether an AI initiative is successful or delivers ROI?

Q3.2: Are there specific frameworks, metrics, or KPIs used to evaluate AI performance?

Q3.3: How are non-financial effects - such as innovation, decision quality, or customer experience - considered in your evaluations?

Q3.4: Which teams are usually involved in the evaluation process; finance, IT, data, strategy?

Q3.5: How important is ROI measurement for AI initiatives in your organization at the moment?

Challenges and Barriers: *Identify difficulties and organizational constraints in evaluating AI.*

Q4.1: What are the main challenges your organization faces when assessing the ROI of AI initiatives?

Q4.2: How do issues like data quality, attribution, or long time horizons affect the measurement process?

Q4.3: How do you handle situations where AI outcomes are uncertain or hard to quantify?

Outlook and Reflection: *Capture forward-looking insights and lessons learned.*

Q5.1: From your perspective, how should companies improve the way they measure and communicate the value of AI?

Q5.2: What trends or changes do you expect in how organizations will evaluate AI success in the future?