VILNIUS UNIVERSITY
FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

STRATEGIC MANAGEMENT OF INFORMATION SYSTEMS

Vilius Aleksynas

Master Thesis

Dirbtinio intelekto diegimas organizacijose: kliūtys ir veiksniai, padedantys pereiti nuo bandomojo projekto prie praktinio taikymo	**Artificial Intelligence Implementation in Organizations: Barriers and enablers for Transition from Pilot to Production**

Academic Supervisor
Assistant Professor Dr. Justina Valentukevičė

Vilnius, 2025

CONTENTS

INTRODUCTION

Relevance and background of the topic. Artificial Intelligence has become a transformational force in contemporary business environments that is fundamentally changing the way organizations drive innovation, gain efficiency, and achieve competitive differentiation. The transformative power of AI lies in changing conventional ways, automating processes, gaining deep insights from large volumes of data, and adapting dynamically to changes in the market (Krishnan & Mariappan, 2024). By embedding AI, organizations can achieve levels of precision, scalability, and speed that were previously unimaginable. The rapid growth of business, along with growing competition in the markets, compels businesses to reconsider the pathway of AI toward these challenges. Examples include enabling predictive analytics that anticipate market trends and customer behaviors, encouraging proactive decision-making. Additionally, AI-driven tools improve the customer experience with personalization at each touchpoint, driving strong customer loyalty in the end. This will also integrate and enforce operational efficiency due to the reduction of human errors and the optimization of the deployment of resources (Keskin, 2020). Although diffusion of AI technology is proceeding very fast, the findings of the most recent evidence suggest that the ability of organizations to create value from the use of this technology is limited. The recent report from MIT NANDA report that, despite the estimated expenditure of $30-40 billion, 95% of the corporations achieve zero ROI, while only 5% of the task-specific and integrated GenAI pilots achieve production thresholds (Challapally et al., 2025). With businesses increasingly adopting AI technologies, their applications and implications are crucial to understand for gaining sustainable success in a globally turbulent marketplace.

Novelty of the topic. AI continuously redefines new frontiers in technological and economic landscapes, capable of things that were deemed impossible just a couple of decades ago. This is not some evaluative shift in the development of traditional computational methods but, instead, a jump toward systems that can learn, adapt, and make decisions on their own. Whereas the earlier phases of AI development were based on rule-based algorithms with restricted computational power, today's AI is enabled by large datasets, deeper neural networks, and significant computational resources that solve problems dynamically in real-world complex situations (Nayak & Walton, 2024). Over the past decade, AI has moved beyond simple automation to include generative models, autonomous systems, and real-time analytics-a paradigmatic shift both in its application and conceptualization. This research, therefore, investigates these emerging trends, the ever-changing role of AI in business, and its increasing importance in shaping contemporary industries.

Scientific problem. Though overall it has been envisioned to be an revolutionary business strategy driver in itself, most businesses have been less capable of realizing tangible returns for their AI efforts. Such gaps typically arise due to unspecified goal definitions at the time of execution, intrapersonal resistances, strategic incompatibilities, and an under-readiness organizational environment. AI therefore tends to be adopted symbolically instead of operationally and therefore remains at arm's length in delivering tangible business value.

Research Object. The object of this research is the process of AI system implementation and adoption in organizations, including both organizational and technical aspects that influence its success or failure.

Research Goal. To find and investigate the enablers, barriers, and conditions that influence the effective strategic integration of artificial intelligence in firms, and to offer practical recommendations to organizations willing to employ AI in a value-creating way.

Research Objectives:

- To analyze existing scientific literature on AI implementation and adoption in organizations, focusing on the technical and organizational factors that influence success or failure.
- To explore the main challenges and enablers that organizations face when implementing AI systems, particularly during the transition from pilot projects to full deployment.
- To investigate the organizational capabilities that support successful AI adoption, including leadership, data readiness, cultural factors, and cross-functional collaboration.
- To collect and analyze expert insights through semi-structured interviews with professionals involved in AI strategy, development, or consulting.
- To propose practical recommendations for companies seeking to increase their AI maturity and maximize the value generated from AI technologies.

Research Methods and Data Sources. This thesis follows a qualitative approach to research, using semi-structured expert interviews to examine the challenges, enablers, and organizational competencies for AI adoption and implementation. The qualitative research method is appropriate due to the exploratory nature of the research subject and the need to gain in-depth information from practitioners with first-hand experience of involvement in AI projects. This research employs a purposive sampling strategy, choosing participants based on the relevance and experience of the participants. The interview participants were sourced from the researcher's network of professional and academic connections in order to guarantee that the participants possess direct experience with AI deployment or strategic adoption programs.

The interview guide is designed based on the themes that emerge from the literature review and contains open questions that solicit actual experience and reflection from the experts. The interviews are conducted online or face-to-face, recorded with the consent of the participants, and transcribed for analysis afterwards. Data to be collected will be analyzed through qualitative content analysis and thematic coding in order to seek recurring patterns and critical success and failure factors across the cases. The goal is to create a grounded understanding of the differences between AI adoption efforts that succeed and others that only reach experimental phases.

Practical and Scientific Significance. This study adds to the expanding body of research into AI and focuses in particular on the implementation and adoption phases - both important but less researched phases in which AI projects tend to fail altogether or bring in tangible value. By understanding the enablers and barriers that organizations experience, the study enhances the picture of the organizational capabilities needed for AI integration success. The research also contributed to the better understanding gap between technical solutions and organizational preparedness. The study develops and adapts extant theory, including that of AI readiness and capability-based approaches, and translates this into the modern business environment.

From a practical perspective, the study gives business managers, IT managers, and consultants valuable insight to help not fall into the AI craze of implementing without bringing value to the company. By the identification of the leading capabilities and barriers (e.g., cultural resistance, quality of the data, strategic alignment deficiency) and enablers (e.g., AI governance, cross-functional teams, leadership sponsorship), the analysis help companies to bypass pitfalls and put effective AI take-up strategies in place. Recommendations based on expert insights can support better decision-making, capacity development, and AI-related planning.

Artificial Intelligence Use Case in the Thesis. AI Tools were primarily used for initial literature research, such as identifying relevant sources, first investigation with summaries (VU

Library Assistant, ChatGPT). As well as for transcription of interview recordings (Google Gemini). Both transcripts and sources were reviewed and verified by the author.

1. THEORETICAL FRAMEWORK AND ANALYSIS OF ARTIFICIAL INTELLIGENCE IMPLEMENTATION IN ORGANIZATIONS

1.1. Evolution and development of Artificial Intelligence

The study of AI has been among the most transformative areas of technological as well as intellectual activity in human history. Though speculation regarding the nature of intelligence as well as the possibility of mechanical intelligence is traced to ancient philosophy (Gençten, 2018), its contemporary concept occurred mainly during the 20th century, based on the advances in mathematics, logic, and computer science.

1.1.1. Early concepts and Theoretical Foundation development

The concept of creating intelligent systems dates back millennia. Early fantasies of artificial beings were captured by Ancient Greek myths, such as that of Talos, a mechanical giant, and the automata of Hero of Alexandria. Philosophers such as Aristotle formally laid the foundations of logical reasoning, hypothesizing formal system/s that would influence computational theories to come. Mechanistic theories of animal behavior were theorized in the 17th century by René Descartes, according to which some kinds of intelligence could be duplicated by machines (Cave & Dihal, 2018).

The 20th century marked the turning point after which speculative ideas began to flow into scientific research. Alan Turing did pioneering work in the 1930s and 1940s by laying a theoretical foundation for computation through the presentation of the Turing Machine as a universal model for algorithmic problem-solving. Turing's question, "Can machines think?" and his proposal of the Turing Test 1950. laid the intellectual foundation for measuring machine intelligence (Russell & Norvig, 2020).

Probably one of the milestone events that firmly established AI as a science was the summer of 1956 at the Dartmouth Summer Research Project on Artificial Intelligence. Organized by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon, it brought together seminal researchers to consider the possibility of machines eventually duplicating human intelligence. It was at this conference that the term "artificial intelligence" was originally coined, along with the vision for the development of systems able to learn, reason, and solve problems on their own, just like human beings. This initial meeting initiated decades of research and development, establishing the conceptual groundwork for AI as both a precise and an ambitious scientific field (Chow, 2021).

Although the preceding section touched on the history and conceptual origins of AI, numerous definitions have been posited throughout the years with implications on their technology paradigms, scopes, and aims. Table 1 below demonstrates this using an overview of selected definitions of artificial intelligence from classical and modern sources, together with their central focuses.

Table 1. Definitions of Artificial Intelligence and Their Key Emphases Across Authors

Author	**Year**	**Definition**	**Key Emphasis**
McCarthy, Minsky, Rochester, Shannon	1956	“Every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.”	Intelligence is computable; goal is to simulate human-like learning and reasoning in machines.
Haenlein, Kaplan	2019	A system’s ability to interpret external data correctly, learn from such data, and use those learnings to achieve specific goals through flexible adaptation.	AI as an adaptive system capable of learning and goal-oriented behavior.
European Commission	2019	Intelligence as in referring to the ability to choose the best action to take in order to achieve the goal under certain criterias.	Integration of perception, decision-making, and action in a feedback loop between environment and system via sensors and actuators.
Russel, Norvig	2020	Rational agent that perceives its environment and acts to achieve its goals.	Divides AI into human-like vs. rational, and thinking vs. acting.

Sources: Adapted from (Chow, 2021; European Commission, 2019; Haenlein & Kaplan, 2019; Russell & Norvig, 2020).

1.1.2. Types of Artificial Intelligence Models

As AI has grown, so have the explanations and categorizations of its scope, operation, and complexity. From the initial focus on symbolic reasoning and the application of rules, contemporary AI encompasses several paradigms of learning as well as capacities to describe systems by both their operation and complexity.

A prevalent method of classifying artificial intelligence is capability, or the scope and generality of the tasks the AI system can execute. Brought to the forefront by (Bostrom, 2014), this categorization describes three broad classes: narrow, general, and superintelligent.

- Artificial Narrow Intelligence : Also referred to as weak AI, such programs are designed to perform only a specific individual task (e.g., recommendation system, spam filtering) and are unable to work beyond their predetermined parameters.
- Artificial General Intelligence: Theoretically existing systems with the ability to match humans across an incredibly wide range of tasks. AGI is theoretical but is being researched actively.
- Artificial Superintelligence: The hypothetical form of AI more advanced than human-level intelligence across all subject areas. Although it is widely debated in strategic foresight and ethics, it is still hypothetical.

Whereas the capability typology is focused on the scope of the tasks the AI system is able to perform, the other leading perspective focuses on the form the AI system works through and processes information. Another such typology, put forward by (Hintze, 2016), categorizes AI along functional dimensions-spanning from reactiveness to hypothetical emergence of self-awareness. This typology is beneficial to an understanding of the development of AI's inner complexity, but also to its potential cognitive development.

- Reactive Machines: This category of machine is able to respond to the current input and has neither memory capacity nor the ability to utilize past experience. The best application of this is the IBM Deep Blue chess program, where moves are taken into account depending upon the current board position.
- Limited Memory: The majority of AI applications belong to this category today. These applications have the capability to use past information to guide their decisions, such as autonomous cars using the history of traffic to guide their decisions on the road.

- Theory of Mind: Higher level and more theoretical type of AI, would be able to comprehend emotions, beliefs, and social cues. These are theoretical as yet, but the idea is a giant step ahead for AI.
- Self-Aware AI: The most speculative and hypothetical category, where the machines are self-aware, possess an understanding of their own inner states, and the capacity to model the inner states of others.

Besides classifying AI on the basis of its capability or functional complexity, the other common perspective is methodological in its conception of constructing the AI systems-that is, the development and method of learning of the AI systems to accomplish the tasks. This classification is between logic-based and data-based learning systems and is of particular significance to consider while making implementation choices, model interpretability, and performance compromises. Symbolic AI, machine learning, deep learning, and reinforcement learning belong to this classification with various legacies of technology development and design tradition (Russell & Norvig, 2020).

- Symbolic AI (Good Old-Fashioned AI): Based on rules, logic, and symbols. They use explicitly stated rules and dominated the early decades of AI development. Expert-systems are the prime example.
- Machine learning: an approach where the system is able to learn to identify patterns of information rather than being directly coded with rules. This approach includes supervised, unsupervised, and semi-supervised learning, and is the basis for much of the application of AI today.
- Deep Learning: This is the application of multilayer neural networks to handle advanced tasks such as the classification of images, natural language processing, or playing video games. GPT and AlphaGo are two examples of this.
- Reinforcement Learning: The agent acquires the best action through trial and error depending on the penalty or the reward. Reinforcement learning is used widely across video games and robotics to model sequential decision-making.

It is worthwhile to understand the ways AI can be classified-by capability, function, or approach-to position its deployment within the context of the organisation. These classifications not only specify the variations on what AI is and what it can do but also have strategic, technical, and organisational implications for the deployment decision. A reactive AI system compared to the use of a deep learning-based deep model requires fundamentally different requirements for underlying structure, integration, and governance. The foregoing overview thus establishes the necessary context for the discussion to be developed further in the following chapters of the barriers, enablers, and determinants of the journey from AI pilot to production-level deployment.

1.2. Applications of Artificial Intelligence in Organizations

With the technology getting progressively advanced, AI adoption in business practice becomes more widespread and advanced too. Where the earlier section has elaborated upon the variability of AI systems according to capability, functionality, and strategy, the current section addresses the actual utilization of those systems across various businesses. AI is being used differently across businesses-depending on their operating needs, availability of the data, risk appetites, and regulatory landscapes. After the evaluation of AI use across the retail to healthcare to finance to manufacturing sectors, this section attempts to identify not only the ways AI is being leveraged, but also the ways industry-specific factors drive the success, scope, and tenure of AI use. From this comparison, we find out what kind of solutions generalize across businesses and where the need is for custom, domain-specific solutions.

1.2.1. Healthcare and Life Sciences

The healthcare sector demonstrates the potential of AI and the subtle limitations of its adoption. AI is being used across diagnostics, surgical planning, drug development, and treatment interaction. These applications primarily involve the use of machine and deep learning techniques being applied to large sets of data such as electronic medical records, medical images, and genomic profiles.

Some of the highest profile applications are medical imaging. Deep learning algorithms on X-rays, MRI, and CT scans now identify tumors, fractures, and disease of the nervous system-often more precisely than medical professionals. Google DeepMind's AI, for instance, provided record-breaking diagnoses for more than 50 eye diseases (Suleyman, 2018). Machine learning algorithms also support predictive analysis for chronic disease care, including predicting diabetic or cardiovascular disease patient readmission to the hospital (Bernal & Mazo, 2022)

.

However, the gap between theory and practice is broad. The most common barrier listed is trust. Patients and clinicians are slow to follow AI advice even when the model is more accurate empirically (Grote & Berens, 2020). The hesitation arises from accountability concerns, the absence of empathy, and reduced explainability. These challenges are further complicated by ethical risks, algorithmic biases and fairness, since AI models tested on unrepresentative samples may produce incorrect outputs or biased outputs for underrepresented groups.

Firms such as Atomwise and Insilico Medicine use neural networks to model the interaction between two molecules and identify new drugs, dramatically shortening the initial stages of drug research. The technology raises intellectual property implications, explainability, and scientific verification (Sequeira & Tsang, 2024). Regulatory agencies lag behind in approving AI results with non-explainability verification tools because most of the platforms are "black boxes."

Excluding diagnostics and drug development, operational efficiency is being optimized through the use of AI. Clinical notes are automatically transcribed and patients summarized using NLP software, and there is workforce planning algorithm scheduling of physicians to prevent physician burnout (HHMGlobal, 2024). There is the concern, nonetheless, of too much dependence on AI eroding the emotional aspect of care. Patients complain of discomfort with chatbots or virtual help interacting with emotionally charged topics (Ayers et al., 2023).

Precision medicine is another promising area. Based on the constellation of the genetic, environmental, and lifestyle determinants, AI is able to generate individualized treatment regimens. An example is the use of deep reinforcement learning to generate adaptive therapy regimens for prostate cancer by University of Oxford and Moffitt Cancer Center scientists (Gallagher, 2024). Even there, patient receptivity is not the same. Patients have been found to be more amenable to following their physician's advice even closer to the truth with AI-informed advice (Chalutz Ben-Gal, 2023).

These instances highlight an underlying contradiction: narrow AI programs are showing strong, domain-specific competency, but their wider adoption is being held back by psychological, regulatory, and ethical factors. Healthcare is arguably the most ready to be disrupted by AI, yet the most unlikely to see large-scale automation. These forces to an extent contribute to the reality of so many promising healthcare AI programs getting stuck at the pilot study stage, never quite proceeding to everyday clinical practice. The healthcare case thus constitutes an appropriate prism to understand why there are so many AI systems, even when they are technically advanced, which never move beyond the pilot stage. The case shows us the ways in which sectoral features-rigorous regulatory requirements, ethical factors, and deep-seated belief in professional judgment, for instance-shape the channels of AI adoption. These are especially near the conclusion of this thesis: identifying the environmental and organizational drivers of productive use of AI technologies under the context of real business settings. Healthcare, compared to the other sectors addressed later, reflects the extent to which professional norms, risk conservatism, and patient trust may become greater barriers to adoption than technological readiness itself.

1.2.2. Retail and Marketing

Retail and marketing are among the most active business areas where AI has progressed beyond the pilot to actual operational use. The industry showcases the ways AI is streamlining back-end retail processes and tailoring front-end consumer interfaces-offering useful comparison to more heavily regulated or slower-moving sectors such as healthcare.

AI technologies support retail demand forecasting, dynamic pricing, and recommendation. Machine-learning algorithms provided with information on past sales, seasonal demand, and external factors such as weather and economic indicators reduce wastage and stockouts. AI is employed to maximize warehouse inventories at Wal Mart, for instance, to increase the availability of goods and reduce wastage (Musani, 2023). The algorithms support the re-routing of goods in real-time to increase speed of delivery and customer satisfaction.

Dynamic pricing is another crucial use where businesses can dynamically change prices, depending on availability, demand, and competitor behavior. Businesses such as Amazon are changing millions of prices on a daily basis using artificial intelligence to stay competitive and generate the highest possible revenue (Morello, 2024). Others are emulating what is being done across such sectors as aviation where dynamic re-pricing of flights is done to respond to dynamic markets (Walker, 2023).

Recommendation engines drive much of the client-side purchasing. Collaborative filtering and content-based approaches underlie the model of Amazon and provide the ultimate personalization through browsing history and purchasing history. These not only drive the conversion but also engage the user more significantly. Additional complementary personalization approaches further refine the outputs to other intention, interest, and geography factors (Alaa, 2024).

AI helps deliver and generate personalized content at scale to the target audience for marketing. Machine learning observes the user behavior, social media, and engagement trends to maximize where, when, and how to deliver ads. Google Ads and Meta's Ad Manager dynamically optimize advertisement placement and bidding to deliver ads to high-probability audiences and reduce the wastage of the campaign (ABM, 2023). Ad creation platforms like Jasper AI and Copy.ai help marketers generate advertisements, emails, and social media posts at high speed (Jandal, 2024), whereas visual-generation platforms like DALL·E help companies generate visual content without feeding through traditional design funnels (PixelFish Media, 2024).

AI has also revolutionized retail customer support. Virtual chat assistants and chatbots with natural-language processing respond to queries, execute orders, and handle most basic problems 24/7. Sephora's chatbot, for example, not only replies to messages but also makes suggestions depending on personal taste (McKinnon, 2018). Advanced sentiment analysis technology such as IBM Watson can even detect frustration and disappointment on the consumer's side in real time so businesses can emphasize service excellence and optimize support accordingly (IBM, 2023).

Though the retail sector has welcomed AI with open arms, there are certain pitfalls too. There is intrinsic bias in recommendation algorithms resulting in echo chambers of circular recommendations, lowering the diversity of discovery. Over-personalization is also an issue on the use of the data and privacy side. AI-powered content ease has ethical concerns related to its authenticity, to say nothing of consumer manipulation.

However, compared to healthcare, retail firms have fewer ethical concerns to grapple with, and therefore experimentation is quicker and scaling is easier. This sector is selected to demonstrate the manner in which relatively low regulatory compliance, high competition, and measurable ROI drive the application of AI. These forces have applicable context for this thesis to examine what makes AI adoption profitable, large-scale, and sustainable across sectors. As latter sections demonstrate, not all sectors can maintain such velocity-making retail the strongest case study of speed, nimbleness, and consumer-centric adoption of AI solutions.

1.2.3. Manufacturing and Logistics

Production and logistics are the promise of AI where efficiency, accuracy, and integration of the system are the most important factors. Adoption of AI is not consumer- and personalize-directed, such as it is for retail, but rather prompted by the need for operating efficiency, predictive maintenance, and smart automation. Healthcare, the other sector where regulatory and ethical hills slow down adoption, has to overcome more technology and integration organizational barriers.

AI is among the key drivers of smart manufacturing where the physical processes get closely integrated with the cyber infrastructures. The integration becomes possible through AI-fused digital twins, offering the real-time duplicates of the running processes and facilitating simulation to support decision-making and enhance efficiency (Tao et al., 2019). The processes get connected to the Industrial Internet of Things (IIoT) to receive real-time sensor feedback for enhancing the processes.

Predictive maintenance is perhaps the most promising application of AI to have been implemented in the production sector. Machine-learning algorithms are engineered from sensor signals to predict the failure of machines prior to it taking place, and this serves to cut back on downtime and the life extension of assets alike (Zhang et al., 2017). The practice is significantly dissimilar to the traditional response or planned maintenance practice and has been used across numerous sectors relying on heavy machinery and constant production.

Meanwhile, logistics processes have been automated using such AI platforms to optimize the delivery time, route, and fuel consumption according to traffic, weather, and road conditions (Atom, 2024). Warehouses also rely upon robotic machinery to get packages, orders, and inventories processed automatically, with solutions such as those offered by Siemens' SIMATIC Robot Pick to handle cumbersome items through the elimination of manual intervention (Siemens, 2023). AI also aids the supply chain's resilience through the use of predictive analytics-used by companies such as Maersk-to predict disruptions caused through geopolitical tensions or sluggish suppliers to enable forward-thinking operational redesigns (Rebello, 2024).

In reality, AI deployment in the context of manufacturing begins with operational proofs-of-concept, such as predictive analytics on individual pieces of equipment or AI-assisted quality inspection on individual production lines. These proofs-of-concept remain isolated, though, due to integration requirements, lack of qualified personnel, or organizational reluctance to digital transformation (Chatterjee et al., 2021). Siemens and Bosch are among the companies to have achieved success integrating AI across end-to-end value chains using modular infrastructures and AI-powered platforms, but such instances remain the exception. Even when ROI is obvious, the leap to companywide AI adoption requires strong digital maturity and synchronization among operations, IT, and C-level functions.

The case of logistics and manufacturing is the central topic of this thesis because it is an area where return on investment is measurable and where the result is immediate but where organisational readiness and integration of the system fall behind to derive maximum potential from the power of AI. As compared to the healthcare context, where there is the factor of ethics to beware of, and the retail context, where there is the need to get it right for the end consumer, the case of manufacturing emphasizes the role played by such factors as technical compatibility, legacy, and digital maturity to the adoption of AI.

1.2.4. Banking and Finance

AI origins run deep in the finance sector today and are changing not only high-risk trading but even back-office functions and everyday banking functions. As with the logistics and the manufacturing business, the journey to AI is facilitated through the need for accuracy, efficiency, and automation. However, compared to the manufacturing business, where the return on investment is realized through the material good, and the healthcare business, where the absorption is slowed down through ethical and legislative concerns, the finance sector is involved in the high-digital, information economy-suited to the applications of AI on large scales.

Among the highest-profile uses of the technology is algorithmic trading, where machine algorithms search history and live information and execute trades instantly. Investment firms and hedge funds use AI to detect micro-trends across markets, trade high-frequency, and maximize portfolio return. Quantitative funds such as Renaissance Technologies use proprietary algorithms to search for arbitrage opportunities and reduce exposure to risk (Verma, 2024). The algorithms are controversial, nonetheless. High-frequency trading has been shown to lead to rising volatility and falling transparency, such that system risk and regulation have become the subject of concern (Bathaee, 2018; Kirilenko et al., 2011).

For credit risk evaluation, AI is increasingly being used to gauge borrowers more comprehensively. Not only their history but their conduct-payment history, social media activity included-machine algorithms are allowing more accurate and inclusive lending approval (Faheem, 2021). With greater predictive capability compared to traditional methods but still causing constant worries for algorithm bias and transparency, to say nothing of when used within regulated settings.

Anomaly detection is yet another where AI has left its mark. AI programs are continuously watching for user action and transaction trends to identify anomalies in real-time. JPMorgan Chase, for instance, uses AI algorithms to track millions of transactions and find the potential for fraud-both cutting down the loss to fraud as well as building consumer confidence. High rates of false positives are, however, an operational thorn, at times causing consumer frustration when legitimate transactions get flagged erroneously (Lucinity, 2024).

Beyond such areas of application, the interaction between banks and customers is being reimagined as well through the use of AI. Virtual assistance and chatbots such as Bank of America's Erica utilize natural language processing to support customers 24/7, offer personalized advice on finance, and assist customers in planning their budget. To date, since 2018, there have been over 1.5 billion interaction instances through the application across 37 million customers, hugely cutting back on the number of calls to the contact centers and the service complexity (Aldridge, 2023). This is similar to retail space areas of application where responsiveness and customization are being optimized through such assistants-albeit with higher compliance requirements in finance. Above all, nevertheless, AI is not limited to client-facing roles. The majority of the finance industry is comprised of office work made up of knowledge-work, and increasingly this is being streamlined through the application of automation. Intelligent document processing technology and Robotic Process Automation (RPA) are now used for regulatory reporting, contract interpretation, and loan origination automations. JPMorgan's COiN solution is used, for instance, where intelligent language processing is used in processing tens of thousands of contract documents in seconds-otherwise taking time for human beings annually of up to 360,000 hours (Borczuk, 2024). These kinds of tools are worth their weight in gold to reduce the risk of human error, speed the workflow, and free staff to focus on more judgmental or more complex work. Though finance is digitally mature compared to healthcare and logistics, its challenges are unique too. These range from stringent privacy regulations such as the GDPR to advanced regulatory requirements around accountability and transparency. The "black box" of most machine algorithms makes explainability, the basis of auditability in finance, more challenging. Institutions therefore find themselves obliged to invest not only in AI solution sets but governance frameworks and risk mitigants to enable compliant deployments (Turner, 2024).

Financials overall represent the promise and the dilemma of AI deployment. With its intersection of structured data, regulatory needs, and dependence on digital processes, finance is midway between the measurable operational efficiencies gained in logistics and production and people-problems challenges encountered in healthcare. The trajectory it is on is not only that of technological development, but of the capacity of institutions to bring AI systems to equilibrium with ethical, regulatory, and organisational equilibria.

1.2.5. Summary of Sectoral Comparisons

The cross-industry overview given indicates the use and scope of AI are skewed despite the promise it has to transform the sectors. Healthcare, despite the vast volumes of available data and potential for innovative applications, is held back by ethical brakes and cultural hesitancy. Retail and e-commerce are the leaders among large-scale rollouts due to obvious ROI, light-touch regulations, and high rivalry. Manufacturing and logistics are held back by high costs of integration and legacy system resistance, albeit their process efficiency gains are tantalizing. Finance, digitally advanced but tightly regulated, represents the promise along with the complexity of AI transformation. These trends would mean technology capability is but part of the driver for AI adoption; sectoral contingencies-regulatory push, culture readiness, organizational form- drive the equation. This is the thesis to study not only the functional promise of AI but the circumstances under which the adoption of its use is feasible, scalable, and sustainable-a thread to hold the chapters together to follow.

1.3. Artificial Intelligence Implementation

While there is a variation in AI implementation understanding between organizations and experts, AI implementation is generally described as developing, deploying, and integrating AI systems within business operations. Or in other words a technical process of coordination of data, infrastructure and system integration (Weber et al., 2022).

1.3.1. Artificial Intelligence Implementation Process

Weber et al., (2022) outline the implementation process as a series of interactive phases, from problem formulation to model development, deployment, and continued monitoring. In contrast, Halper, (2025) paints a more business-oriented picture, with a focus on steps such as business understanding, modelling , assessment, and maintenance after deployment. Though different in emphasis, both support the same structure that organizations need to follow as best practices. Combination of both views might be a good fit for organizations to start AI development, visible in Table 2.

Although the steps are a natural progression, AI implementation in the real world is never linear. Organizations will oftentimes go through prior steps due to new information, shifting objectives, or performance variation in the models. Iteration among data preparation, modeling, and testing is especially prevalent. Furthermore, the rapid pace of change of AI creates a dynamic setting: unlike traditional information systems, which are relatively stable for years, AI platforms and tools are outdated in a matter of months. Such rapidity requires agile processes, tight feedback loops, and cross-functional collaboration.

For all of these adjustments, though, most organizations still suffer basic issues in their deployments, integration, and post-deployment surveillance-that is, when AI systems must interact with physical-world infrastructure and working patterns. Late-stage failures are associated with low-level issues with the data: poor-quality information, insufficient engineering talent, and skewed datasets erode performance in production (Ryseff et al., 2024). Meanwhile (Bojinov, 2023) outlines organizational weaknesses, for instance, unclear project ownership, unrealistic assumptions, and poor alignment among business and tech staff. While organizational issues following these will be addressed more in Section 1.4, it is important to recognize here that most technical failures are symptoms of broader strategic and structural weaknesses.

Table 2. Artificial Intelligence implementation stages

No.	Stage	Description
1.	Business Understanding	Clarify the business problems, inefficiencies, strategic goals and look over or define key performance indicators.
2.	Data Collection and Preparation	Gather clean and fitting data in order to avoid “garbage in garbage out”. A lot of the time vectoring of data is required for AI model training.
3.	Model Development	Choose a good model, algorithm or a technology, train them on the data, validate initial performance.
4.	Model Evaluation and Validation	Access model accuracy, correctness and robustness depending on the business needs, and check performance costs.
5.	Deployment and Integration	Move the validated model into production and incorporate it into the business workflows.
6.	Monitoring and Maintenance	Continuously track model performance, retrain or adjust as performance changes, also keep track of new requirements or errors.

Sources: Adapted from Weber et al., (2022) and Halper, (2025)

1.3.2. Technical Challenges in Artificial Intelligence Implementation

While the above did describe a sequential implementation process-from business comprehension through to deployment and upkeep-a pragmatic execution of the steps rarely occurs in a straight line. Each step in the implementation process has a portion of technical issues, and even those that have a well-laid road map still experience unplanned resistance. Organizational issues are a substantial contributor (as discussed in section 1.4.) though often the reasons for most implementation slowdowns are technical in origin and occur considerably earlier than we are wont to think.

A typical issue is the disparity between AI as a notion and the specific issues to which it is applied. The solution-first path for AI, rather than as a set of niche methods, leads most businesses to initiate implementation without ensuring whether AI is an appropriate solution for the problem (Ajuzieogu, 2024). These projects do not succeed in those cases since they don’t have clearly defined problem statements, domain expertise, or apply overly advanced models for relatively simple business needs.

Data quality and readiness is arguably the most cited technical challenge. Being normally a tech asset, in the field, it is linked to organizational processes, governance controls, and legacy systems. In (Ajuzieogu, 2024), a staggering 68% of AI project failures are caused by issues in the quality of the data, i.e., inconsistency, incompleteness, or bias. Such issues are typically prevalent in unstructured inputs (i.e., image, audio, or natural language input), where preprocessing and labeling are extremely labor-intensive and subjective.

Also inhibiting implementation is the technical impediment of integration into existing systems. Models, in theory, are deployed with ease and tested. Reality sets in, and integration obstructions arise when models require integration with enterprise resource planning applications, legacy applications, or multi-vendor, heterogeneous software stacks between departments. Both (Halper, 2025) and (Weber et al., 2022) report this process is most commonly underestimated during planning. Flaws in pipeline design, incompatibility through APIs, or latency constraints in near-real-time may all impede functional deployment regardless of the performance of a model in isolation.

Even when implemented technically, explainability of models is still a challenge, especially in applications where traceability and user trust are necessary. Ensemble methods and deep learning are especially "black boxes," producing outputs for which reasons and mechanisms may not be straightforward to explain. In applications with a high stake, for which finance and health care are examples, lack of transparency leads to legal noncompliance or damage to reputation. Explainability is not an aspirational requested feature, but rather a necessity, one that, if overlooked, is a strong barrier to scale and deployment (Arrieta et al., 2019).

A strong example of this threat is the case of Target's application of predictive analytics, which predicted a teen shopper's pregnancy before she and her family were aware and sent her marketing messages. Operationally effective, the incident revealed social, ethical, and reputational dangers of uninterpretable and ungoverned AI (Piatetsky, 2014).

Last, yet most often neglected, is post-deployment monitoring. AI models are not fixed instruments-they degrade over time due to data drift, new business needs, or ambient variation. In the absence of deliberate retraining, versioning, and anomaly detection, businesses rely on increasingly inexact and brittle systems over time. In most cases, businesses assume execution of a working solution is where technical work is done, when it is in fact the beginning of long-term operational stewardship (Klaise et al., 2020).

These technology issues-from the quality of the data, through explainability of models, to integration into systems-are not implementation-dependent issues. They are rather ubiquitous barriers to AI projects' scaling from pilots to broader organizational use, a central problem this thesis examines. By exposing the character and impact of such issues, this work aims not only to describe how things go wrong, but under which technical and strategic circumstances AI systems are able to move from experimentation into value-delivering deployments. These results will also correlate with organizational enablers and capabilities elaborated later in Sections 1.4 and 1.5, and inform expert opinions investigated in the thesis's empirical chapter.

1.3.3. Technical Enablers for Successful Artificial Intelligence Implementation

While technical hurdles increasingly have the spotlight in discussions surrounding AI deployment, high-performing organizations increasingly reflect a set of enabling capabilities and practices. Instead of merely solving immediate issues, these in fact prevent the likelihood of project failure by embedding resilience, flexibility, and transparency within implementation.

The root enabler is the development of mature data infrastructure. This is greater than data correctness - it also includes data pipeline automation, versioning, tracking lineage, and metadata governance. Firms that invest yearly in data architecture (i.e. Data lakes, ETL/ELT pipelines, and master data management) build the foundation required for reliable model input and continuous learning. A single platform for data and clearly established data domain ownership enable AI teams to work more efficiently and avoid common pitfalls such as siloed inputs or schema drift.

Practiced closely in tandem is the use of Machine Learning Operations (MLOps) introduced (Rajuhegde, 2024). These include continuous integration and deployment (CI/CD) of models, automated testing, monitoring of performance, and retraining pipelines. It is argued that MLOps is a key differentiator of companies that are always succeeding with AI, and others stuck in pilot purgatory. Implemented properly, MLOps not only facilitates deployment, but lifecycle model management - making it possible for organizations to track changes, detect performance degradation, and respond ahead of time (Weber et al., 2022)

One of the primary technical enablers is modular and API-first architecture. Rather than building AI components as self-contained or monolithic offerings, top-performing teams focus on interoperability and scalability. That means containerization (Docker), microservices, and RESTful APIs to enable models to be pluggable into existing systems, dashboards, or other external apps without profound rewrites. Modular architecture also enables portability across cloud and on-prem environments - a growing requirement in hybrid or regulated IT stack spaces.

Further, explainability and transparency tooling is no longer seen as a cost but a technical necessity. Organizations that incorporate interpretability into model development - perhaps through SHAP, LIME, or attention mechanisms - will enhance regulatory compliance and build user trust. Explainability tooling is being integrated into development platforms and environments to help developers see decisions, indicate anomalies, and communicate with stakeholders about model reasoning.

Finally, all these technology enablers perform best if they integrate into clearly established business workflows and objectives. Not only should AI systems generate prediction, but they should also be fed into decision-making and action. Successful companies are more likely to infuse AI capabilities into existing tools - CRM systems, ERP dashboards, say - so the predictions can be interpreted and acted upon within the natural flow of work. There is a necessity of owning a business and not using a model, so that AI is not technically in a vacuum but for quantifiable strategic reasons (Weber et al., 2022).

While a similar set of enablers appears in practitioner models as well as in literature, the practice of their deployment, particularly in resource or time-constrained settings, is variably reported. It is unclear how companies assign, sequence, or tailor enablers to particular industry contexts or to various phases of firm maturity. These are questions literature is unable to provide an answer to, but which this thesis will try to provide through extensive expert interviews.

1.3.4. Different Methods of Artificial Intelligence Implementations

While the following sections described phases, barriers, and enablers of AI adoption, they necessarily depict one single progression. In practice, though, businesses address AI integration through a combination of approaches-that is, custom development, off-the-shelf, or hybrid approaches. All three approaches have varying implications for tech complexity, integration efforts, and ownership by the organization. Knowledge of such implementation modalities is thus worthwhile in order to frame the following barriers and enablers in context. Moreover, such decisions are the bases for actual trade-offs elaborated in expert interviews in this thesis.

The implementation of AI into modern businesses represents a major step toward unlocking its full potential. Pre-trained models such as OpenAI's GPT or Google's BERT have greatly simplified AI adoption, enabling organizations to harness advanced capabilities without the complexity of building and training systems from scratch. These models reduce the technical barrier by providing ready-to-use frameworks that smoothen deployment with flexible options, making integration more accessible and efficient for businesses across various sectors (OneAI, 2023).

1.3.4.1. Custom Built Artificial Intelligence Systems

Custom or bespoke AI solutions are developed from scratch for specific organizational or sector needs. They are more suited to in-house processes than off-the-shelf and, in most situations, are the only option where sector-specific practices and regulation require customized logic and functionality. Citing (Liul, 2022), they offer complete system design, algorithm selection, and control by default on the data, resulting in a best-fit solution in extremely sensitive areas such as finance or pharmaceuticals, where confidentiality and adherence are of the highest value.

Among the benefits of tailor-made AI is the fact that it is able to be customized. Such models are able to be domain- and performance-tuned to degrees not possible for pre-trained or off-the-shelf models. A bank, for example, will be able to craft an engine for detecting fraud trained on client-specific patterns of activity, or a pharmaceutical company will be able to calibrate new molecular behavior in a process of developing a new medicine-something pre-trained applications are incapable of quantifying in any substantial way.

In addition, custom development leads to innovations. One can experiment with reinforcement learning or a generative adversarial network in test environments that may not be feasible in pre-built systems. (Bhartiya, 2024) describes this advantage in complicated manufacturing environments, where AI would need to operate in conjunction with expert hardware, inputs, and working environments. Advantages, though, are obtained at a steep cost. (Nguyen, 2024) advises that tailor-made AI involves heavy investment of talent, time, and resources in engineering, in data science, and in domain knowledge. Development is long and cyclical, further considering the nature of the dynamic evolution of both AI technologies and of company needs. And it is not a one-time operation: the systems have to be constantly reupdated, retrained, and technologically supported in order to be responsive to evolving working conditions. Literature is correct in that tailor-made AI provides highest strategic alignment and highest implementation effort. Such a trade-off raises a fundamental question for research in the domain of feasibility, i.e., in what types of organizations is this investment viable in terms of maturity, organizational structure, and scale trade-offs. These will be handled in more detail in the empirical part of the thesis, i.e., in the form of expert opinions on ambition, investment, and scale trade-offs.

1.3.4.2. Off the Shelf Artificial Intelligence Solutions

Pre-built AI solutions are pre-developed commercial platforms and tools from third-party vendors for dealing with common business needs. They are pre-trained models that offer rapid-deployment and low entry barriers and enable organizations to leverage AI capabilities without domain-specific in-house technical expertise. (Han et al., 2021) say that this is the most suited for those organizations that are looking for rapid realization of wins from the areas of automation, analytics, or customer experience.

The main advantage of off-the-shelf AI is ease of adoption. Natural-language interfaces, cloud-based integration, and preconfigured functionality-in the case of Salesforce Einstein or Microsoft Azure AI, for example-are all designed to allow organizations to install AI in processes with minimal adaptation. Economic viability is a further advantage. By spreading development and infrastructure costs over a large number of customers, suppliers offer cutting-edge AI functionality at a fraction of the cost of custom development, and are readily accessible to small and medium-sized enterprises.

In particular, off-the-shelf tools will have long-term vendor support, regular updates, and performance improvements fueled by the large number of users. The collective learning cycle among the adopters benefits from the vendor's development, and they do not need to invest further in their behalf, providing continued evolution and optimization of the features (Han et al., 2021).

However, this broad usage also puts a limitation on the flexibility of the tools. Most of the models are tuned for typical cases and may work suboptimally for domain-specific or nonstandard cases. Organisations may need to tailor their processes to accommodate the tool rather than vice versa-sacrificing process effectiveness or competitiveness. (Lin, 2025) refers that this incompatibility is strongest within highly regulated or specialty domains. There are also structural issues. Such over-reliance on third parties raises issues related to data security, privacy, and compliance. The threat of vendor lock-in, through which businesses cannot switch to different platforms or bring in outside tools, can strangle long-term agility and innovation.

Despite its limitations, off-the-shelf AI is still highly relevant for standardized applications such as chatbots, marketing automation, and predictive customer behavior. Some of the examples of AI-driven decision-making that are effectively being plugged into existing systems with minimal friction are those of HubSpot, Zendesk, and Tableau.

Convenient, plug-and-play AI provides rapid and low-cost entry to artificial intelligence, albeit sometimes at the expense of strategic leverage and tailor-made capabilities. As will be described in this thesis, organisations have to weigh ease of solution against whether in the longer term they will produce for value creation-a challenge experts will frame in the language of project boundaries, industry maturity, and AI literacy in the company.

1.3.4.3. Hybrid Approaches

Hybrid solutions establish a middle ground strategically between off-the-shelf and customized AI solutions. These systems start with mature platforms and append the necessary customization for domain-specific needs. In contrast to having to choose between ease and flexibility, hybrid systems attempt to find some middle ground for both and offer faster deployment than completely customized models without losing the flexibility for more advanced applications.

The most important advantage about hybrid models is that they are configurable. They allow you to begin with a tried and tested foundation-a cloud-based recommendation engine or a language model, for instance-and to customize individual parts selectively. That reduces time-to-value and cost of development, respectively, and allows in-house organizations to invest in areas of the system that offer more competitive differentiation. (Morgan, 2022) said the approach is best suited for dynamic markets where business needs change more quickly than full-cycle development plans are able to support.

Also, hybrid AI enables incremental scaling. As organizational needs evolve, organizations are able to scale or deepen the AI capability incrementally-adapting extra pieces at a later date, developing custom data pipelines, or integrating the system more comprehensively into workflows. (Marr, 2024) suggests that hybrid solutions enable a phased path for moving from a proof-of-concept to enterprise implementation without having to go on a complete build from scratch.

This flexibility also provides hybrid AI with particular value in regulated environments. Delicate functions remain inhouse through customized modules, while less sensitive parts-such as third-party APIs or end-user interfaces-are outsourced to well-known providers. Both parts of this setup can reconcile information sovereignty and compliance issues of concern with efficiency in operation. All that said, hybrid designs are not free from challenges. Integration complexity increases geometrically as disparate pieces-from proprietary and vendor-controlled ownership-that need to work together reliably. Compatibility among tools and frameworks needs to be addressed by organizations, which in some cases involves spending on middleware, APIs, and cross-functional planning. Maintenance becomes more complicated with shared responsibility between internal organizations and third parties, and strong governance is necessary to facilitate system dependability and accountability.

Hybrid approaches are becoming popular where domain-specific accuracy and broad applicability are both desired. In medicine, for example, hospitals are able to marry an off-the-shelf electronic health record (EHR) system with a locally trained customized diagnostic imaging model for the local patient population. In logistics, businesses are using a vendor-supplied routing optimization module augmented by locally developed in-house proprietary supply chain risk forecast models.

Hybrid AI deployments are thus a pragmatic solution for those organizations that require velocity and control. Their success, in conception here through expert interviews, appears to rely on an enterprise's technology maturity, its ability to successfully integrate offerings, and its collaboration between and among in- and out-side stakeholders. More broadly still, the choice among off-the-shelf, custom-built, and hybrid AI carries further trade-offs among scalability, flexibility, and complexity. With more hype encumbering AI, too often projects still fail at pilots not by outright technical failures, but by mismatches between implementation styles and organizational preparedness. How practitioners make those trade-offs is critical to observing and learning from paths to sustainable AI adoption.

1.4. Artificial Intelligence Adoption Organizational Readiness

AI is also an organisational transformation problem and not only a technological problem. If organisations are headed towards end-to-end, production-grade implementation of AI, it is obvious enough that the largest barriers are not technological - but internal, and located in culture, structure, capacity, and trust. To succeed with AI, it is not about having the right tools and data, but employees, legal environment, and organisational culture well aligned and enabled to leverage them.

Several authors have used various labels to categorize the factors of AI readiness and adoption. (Halper, 2025) considers people, process, technology, and governance as being of concern. (Weber et al., 2022) classify readiness under leadership, skills, and implementation challenges, while the (OECD, 2022) considers legal protection, responsibility, and public trust as areas of focus. These definitions agree on the need for readiness to reach far beyond infrastructure This research builds on these considerations, focusing on four core areas of readiness to pursue AI: organizational, legal, ethical, and social. They weren't chosen simply because literature is full of them, but because they kept re-emerging in early research and practitioner consultation. Of special note, this chapter addresses talent and capability gaps - one of the under-emphasized but critical components of the adoption environment of an enterprise.

1.4.1. Organizational Challenges

Not only technical challenges arise when trying to implement AI into business processes. In fact a big portion of challenges are actually organizational - depending on the existing structure, governance , and cultural preparedness of a firm to undertake adoption of such disrupting technology. AI adoption integrates and reassembles conventional processes, disrupts traditional control of decision-making, and necessitates cross-functional collaboration that many organizations are not yet prepared to direct.

Organizational readiness or in other words, to what extent the company's current processes, infrastructure, and staff are aligned to AI deployment and development is one of the biggest challenges. Companies tend to underestimate the difficulty of aligning operational goals, end-user behavior, and current IT with AI systems. Misalignment puts adoption in deep freeze or creates disappointing results that dispel internal enthusiasm with AI initiatives in the first place. (Weber et al., 2022)

Another often overlooked issue is the lack of clearly defined AI strategy. Firms may propel AI as a buzzword, rather than a business driver, leading to siloed pilot projects with unclear ownership or performance metrics. Lacking leadership support and well-defined goals, technologically well-improved AI models may fail to be adopted or get incorporated into routine decision-making.(Halper, 2025)

Change management is likely the most underrated yet critical organizational challenge in successful AI adoption. The introduction of AI typically brings not just new hardware, but a revolution in responsibilities, procedures, and decision-making models. Deployment of AI requires an intentional shift in how people work, what they value in terms of performance, and decisions-changes that can erode accepted mores and cause resistance if confronted reactively. Staff may be wary of or even hostile toward AI solutions for various reasons: loss of job due to automation, skepticism about the accuracy or fairness of algorithmic judgment, or infringement on control of their decisions. These sentiments are especially experienced in organizations lacking high digital competence, open communication, or openness of culture.(Wrenn & Sohn, 2021).

To combat such challenges, effective change management needs to be embedded within the AI adoption phase right from the beginning. This entails initial and continuous engagement of the stakeholders, proper communication regarding the role and ability of AI, and specific training sessions that de-mystify the AI systems and enable the employees to operate with them. As Booz Allen puts it, openness and support-based trust is what will drive success in the long run-and without that, even technologically precise AI deployment can fall victim to insider resistance or flat-out rejection.

1.4.2. Legal and Regulatory Issues

Legal complications abound in the adoption of AI, especially regarding data privacy, security, and accountability. Regulations like the GDPR in Europe, CCPA in the United States, and emerging global standards put strict guidelines on how businesses can collect, process, and store data. Non-compliance can result in serious financial penalties besides reputational damage-a fact that makes adherence to these frameworks a critical priority for organizations (GDPR advisor, 2024).Apart from this, another essential challenge is that the regulations are not uniform in various jurisdictions. This is a problem for large multinational companies with the compliance of various countries' laws. Still, the pace of development of AI is so fast that corresponding legal frameworks fall behind and thus create loopholes and ambiguities (Wheeler, 2023).

Accountability of AI systems also brings about important legal issues. Liabilities, when decisions by AI go wrong in cases of a financial loss or injury to a person, always pose a hard time in pinning on someone. When, for instance, an accident strikes with an autonomous vehicle, where does liability lie: at the manufacturer, at the software developer, or with the user? These are what must legally be sorted for clarification of various roles and responsibilities of different stakeholders along the AI life cycle (Galkin, 2024).

Another point of controversy is the IP rights with regard to AI-created content and inventions. In most jurisdictions, the question of who owns AI-driven innovations-whether it is the creator of the AI, the user, or the AI itself-is still legally unsettled. For example, an invention of the AI system named DABUS was refused patent protection in many countries, including the United States and Europe, because current IP laws do not recognize non-human entities as inventors (Smith, 2022). In such a way, it revealed the deficiency of the existing frameworks to cover such basic aspects as the role of AI in the innovative process and made some fresh legendaries relevant for the legal definition of such progress. These are fundamental issues with regard to stimulating innovation while giving due credit and compensation to those who innovate.

Also, except for this fact, the concern about surveillance and misuse of AI with its use in face recognition or predictive policing strongly suggests a design of such regulation within tight framing so that it does not overboard itself, compromising civil liberty. These many a time has complaints relating to invasion of private space, which disproportionately has targeted the poor people, being against the whole ethics oversight also under legal framing (Almeida et al., 2022).

Regulatory challenges need to be overcome by the coordination of policymakers, industry leaders, and researchers in setting clear, adaptive, and enforceable guidelines. The best legal frameworks protect not only against risks but also nurture trust and innovation in the process of both developing and deploying AI.

1.4.3. Ethical Concerns

All of these AI workforce, societal, and ethical impacts are interconnected and pose extremely serious concerns. The most fundamental ethics questions for AI are all about bias, equity, transparency, and exploitation. Unbalanced information would produce maddening disparities in hiring programs, loan programs, or policing programs, for instance. Admittedly, most AI models are "black box" in nature, and this inherently raises questions of transparency and explainability because many stakeholders do not understand how decisions are arrived at, thus eroding trust and accountability. Misuse of AI, including surveillance or creating deepfakes, involves great risks for society, placing a high hand on the need for ethical frameworks that make sure of inclusion, accountability, and the prevention of harm (Sheludko, 2023).

The AI capabilities related to automation raise serious workforce displacement and job loss issues in especially those industries which have a lot of repetitive tasks. While AI has created new roles, it is a transition that requires massive efforts toward reskilling and upskilling. Organisations should actively meet this challenge or be subjected to negative socioeconomic impacts and fair growth otherwise. This is further triggered by public concern for misuse of surveillance, self-regulation, and AI for ethics, discouraging further public adoption(Cheong, 2024).

Besides, the wide application of AI is likely to increase inequalities between organizations and countries, depending on different levels of access to AI technologies. Weaker nations will be losing out competing with AI-powered economies and could create increasing digital divisions between countries. In solving these far-reaching issues, AI expenditure must be accompanied with manpower transformation initiatives and forward-thinking stakeholder engagement for communication with the aim of generating trust and appropriate benefits realization in society (Lumenalta, 2024).

1.4.4. Talent Shortage

While businesses focus on AI initiatives, a major barrier becomes talent shortages. According to (IBM, 2024), as global investments in AI are to reach more than $550 billion, more than 50% of businesses include talent shortage in the limiting factors. The shortage is not only in data scientists but also in machine learning engineers, MLOps professionals, AI designers, and even technical program managers. But the issue lies not only in quantities, but also in alignment.

A report by (Yosifova, 2024) finds a skill mismatch: while hiring managers are looking first for Python, machine learning, and NLP skills, the candidates lack business skills-transcending skills or cross-functional collaboration experience as well. It slows down the implementation of AI, as the models are technically sound but operationally misaligned, as (Deloitte, 2024) found to be a persistent weak area when talent is being used in isolation.

In reply, organizations are turning to internal reskilling to fill gaps rather than solely relying on external hiring. According to (Loh et al., 2024), organizations most likely to expand AI programs are investing strategically in technical and AI literacy training within business functions. The method is distinct from other sectors like banking, where talent is more difficult to build internally. According to The Banker (Crowley, 2025), banks are hampered by chronic deficits of special data-science talent-leading to delayed implementations of AI even with strong digital infrastructure The role of talent barriers is not only in hiring numbers, but in their downstream consequences to the operationalization of AI. Without internal talent to train, incorporate, or oversee AI models, technically successful pilots never reach production. The thesis explores how practitioners actually articulate and fill these talent gaps - not only hiring, but organizational learning, cross-functional coordination, and workforce adaptation as well. Across contrasting methods within different industries, this study aims to ascertain whether talent strategy is an underappreciated determinant of success in scaling AI projects

1.4.5. Organizational Enablers for Successful Artificial Intelligence Adoption

Beyond technical infrastructure, organizational ability to drive implementation is the most critical driver of whether implementation efforts die or thrive. The literature increasingly requires several internal enablers to facilitate transitions between standalone pilots and scalable, large-scale solutions involving AI.

Executive sponsorship is one of the most oft-quoted drivers, according to (McKinsey, 2022). It is closely tied to successful outcomes with AI, not only because it frees fiscal and human capital, but also because it aligns business goals with long-term visions. Projects are prioritized strategically, are made more transparent, and credibility is established across departments, and sponsorship by leadership can bail departments out of organizational paralysis or piecemeal execution.

Executive sponsorship Skills such as cross-functional collaboration as well as AI literacy are another enabler at the center. Organisations continue to be faced with skills gaps-not technical, but business units and non-technical understanding of AI, writes (IBM, 2024). Ground-level understanding of what can and cannot be done by AI is needed, or employees may resist attempted implementation, misuse resultant output, or fail to identify potential applications. The need for agile, multi-disciplinary teams to drive iteration on AI applications in ever-shifting conditions is described by (McKinsey, 2022). Such teams escape the risk of being technically correct yet operationally irrelevant models.

There is a third element to organisational readiness, and that is data governance maturity. Scalable AI is based on trusted, same-time availability of well-governed and well-organised data assets, according to (Halper, 2025). Those organisations with lack of stewardship, controls, or quality assurances are destined to incur build times on their models and constraints on generalisability. Those who do have capability in their systems, though, are well placed to reuse data across applications, retrain at low cost, and respond to changing regulatory needs.

These enablers do not occur in isolation. Each one of them tends to affect the remainder, either reinforcing or inhibiting roles in advances in AI. Empirically, this thesis will look at how the enablers are ranked and viewed by practitioners, and how the enablers affect one another to determine the success of the application of AI in real organizational practices.

1.4.6. Artificial Intelligence Adoption maturity levels

Understanding why AI initiatives fail or succeed cannot be separated from the broader question of organizational AI maturity. As organizations attempt to transition from experimentation to full deployment, maturity models offer a structured way to assess readiness, identify gaps, and guide capability development. These models define stages of progression-typically from ad hoc experimentation to enterprise-wide integration-based on dimensions such as strategy, technology, talent, governance, and culture.

One of the most known frameworks is the Gartner AI Maturity Model, which suggests AI progression in five levels: Awareness, Active, Operational, Systematic and Transformational. At first organizations explore AI opportunities and run pilots with limited business alignment, then AI becomes integrated into core processes, governed by cross-functional teams and tied to measurable strategic outcomes. Gartner's model emphasizes that maturity is not only about having good algorithms, but about embedding AI across multiple teams, decision makers and infrastructure (Semsarpour, 2023).

Similarly, Deloitte's AI Maturity Framework postulates six dimensions-strategy, talent, culture, governance, data, and technology-and positions organizations on four levels: starters, pathseekers, transformers, and leaders. Deloitte estimates very few companies reach "leader" status, with integrated AI governance, domain-based models, and iterated deployment cycles facilitated by MLOps and agile teams (Deloitte, 2022). The majority of companies are stuck in the "pathseeker" stage, with stand-alone applications but scale stifled by lack of vision, lack of coordination, or depth of capabilities.

McKinsey AI capability model focuses on operationalization and cross-functionality. It identifies the fact that high-performing companies build not only technical capabilities but also soft capabilities like experimentation, management of change, as well as business buy-in. Mature organizations, (McKinsey, 2024) finds, are much more likely to see ROI on AI as well as scale initiatives across multiple business units

While such frameworks are valuable diagnostic tools, they are normative and top-down-prescribing idealized, linear progress toward maturity. Reality, of course, is nonlinear, context-dependent, and cyclical. Organizations can be mature in one area (for instance, data infrastructure) and immature in another area (for instance, governance or user uptake). And maturity models must reflect the consultants' and tech vendors' perspective, whereas practitioners are operating under organizational constraints.

From this, one significant space remains unaddressed: how do companies define and construct their own AI maturity? What trade-offs do companies have to make in prioritizing certain dimensions (e.g., velocity over governance)? And how do forces outside the company-industry regulation, talent availability, or strategic imperatives-nudge them toward or away from maturity?

1.5. Why Artificial Intelligence Projects Fail to Scale and lift Pilot to Production

Despite greater investment and buzz around Artificial Intelligence, a number of projects fail to progress beyond pilot projects. (VentureBeat, 2019) cites the fact that nearly 87% of projects never transition to live production environments. These projects always get off on the right foot-displaying excellent proofs of concept in isolated systems-but falter when integrating into real organizational systems becomes involved.

This, termed the “prototype trap,” is one manifestation of a deeper organizational readiness-technical progress misalignment. As already described in Sections 1.3 and 1.4, technically sound projects regarding AI are started by innovation labs or technical departments without any business alignment. Without an understanding of what are the metrics of success, buy-in by stakeholders, or integration planning, technically sound prototypes are never realized as operational worth (Bojinov, 2023; Halper, 2025). When approached as experimentation rather than as a business enabler, scalability suffers.

One of the most typical explanations of why AI does not scale is prototype fragility. The models are, in actuality, developed under idealized sandboxes with clean-labeled data and optimized procedures. The sandboxes are dissimilar to the realities of the actual operating environment, such as legacy infrastructure, heterogeneous data sources, latency constraints, and process variability (VentureBeat, 2019; Weber et al., 2022). When deployed outside controlled lab conditions, the models encounter integration friction, system incompatibilities, or are incapable of producing real-time performance under operating loads.

In the meantime, organizational maturity lags behind technological ambition. As (Chatterjee et al., 2021) observe, especially within sectors such as manufacturing and supply chains, the potential of AI is hampered by legacy infrastructure, data silos, and low digital literacy levels. Even when the model is technically better, it is welcomed with distrust, bewilderment, or resistance on the users' side. As long as workers don't see and trust what can be done with AI-or until business operations are revamped to incorporate them-adoption lags, irrespective of algorithmic effectiveness (Arrieta et al., 2019; IBM, 2024).

This same discrepancy reinforces more profound cultural as well as strategic concerns. Organizations overestimate the amount of change management to scale pilots or one-off projects to enterprise-wide roll-out. As explained in earlier sections, the deployment of AI is widely faced with entrenched norms, roles, and sources of power. Without adequate supporting mechanisms, communication, and reskilling, teams can see AI as threatening autonomy or employment, leading to passive opposition or rejection.

From a governance perspective, most projects also disregard lifespan responsibility, as well. There are no statistics in the context of AI models-they degrade over time as input, business needs, and external conditions change. Without constant tracking, retraining, and lifecycle management, what was good enough in prototype form becomes obsolete or potentially dangerous in live situations(Klaise et al., 2020). The majority of organizations, instead, behave as though deployment is a destination, not the beginning of operational stewardship.

Taken together, these issues suggest that scaling AI is not a technical challenge alone, but a complex, multi-dimensional transformation. Scaling demands:

- Strategic alignment between business goals and AI use cases,
- Executive sponsorship and cross-functional collaboration to secure resources and break silos,
- Robust data infrastructure and lifecycle governance to support system reliability,
- And above all, organizational readiness and cultural adaptability to absorb and sustain change.

While plenty of literature identifies such factors, little literature exists on how organizations are actually implementing them in practice. Existing lists of enablers or levels of maturity in prevailing standards are accompanied by little information on how trade-offs are being made in practice, especially in conditions of time pressure, legacy systems, or low internal capability. This need drives the empirical component of this thesis. By means of semi-structured expert interviewing, the research seeks to study how practitioners across various sectors handle the challenge of scaling AI-prototyping towards production. It is not only interested in understanding why AI projects fail, but also how others succeed, what are the crucial decisions or abilities that differentiate the two, and how scaling is possible sustainably under actual conditions. To this end, the research seeks to build a picture of AI maturity in context-not as an abstract blueprint, but as trial and error, adaptation, and organizational transformation experience.

2. RESEARCH METHODOLOGY ON BARRIERS AND ENABLERS OF ARTIFICIAL INTELLIGENCE IMPLEMENTATION IN ORGANIZATIONS

2.1. Research Aim and Methodological Approach

The aim of this research is to determine the key challenges and drivers of success in adopting AI in organizations, including overcoming the barrier of transitioning from pilots to broad-based operating adoption. It addresses the organizational, technical, and strategic factors of readiness that affect scaling AI initiatives in real-world environments.

Because of the multi-dimensional, context-bound nature of AI adoption, a qualitative, exploratory approach is used in this thesis. It is most suitable in eliciting finer insight into why and how particular organizational events are brought about-most notably in issues where current quantitative models cannot give an adequate understanding of the human, structural, or cultural issues at hand. Also an interpretivism philosophy will be used for this research as to emphasize and better understand different perspectives and challenges in organizations.AI adoption is no longer a technical issue in this scenario but a process of organizational change eliciting professionals to reflect and interpret.

Rather than hypothesis-testing or measuring pre-specified variables, research is guided by the following research questions:

- What organizational and technical barriers hinder the transition of AI projects from pilot to production within companies?
- What factors enable successful scaling and operational integration of AI technologies?
- How do experts across industries assess organizational AI maturity and readiness for large-scale adoption?
- How is AI project success and profitability evaluated?

To explore real AI adoption complexities in depth, this research adopts a practice-based and experience-oriented research path. Semi-structured expert interviews were the methodology of choice in terms of data collection since this allows consideration of actual implementation of practice, issues, and decisions. This is the best approach to uncover nuanced understandings that cannot be ascertained through quantitative or surface analysis.

The interview structure provides the practitioners an opportunity to comment on experience, define successful and unsuccessful efforts, and provide an insight into organizational structure, organizational culture, and strategic alignment and how these factors come to bear on AI initiative scaling. The study is focused on achieving substantive variation among the cases and is more concerned with eliciting the common dynamics and tensions of scaling AI from manufacturing scale up to piloting scale.

The same qualitative research methods have been used in past studies (Halper, 2025; Weber et al., 2022), they investigated readiness of information, organisational AI maturation, governance. Building on this, this research narrows its focus further to the scaling gap: the fact that organizations are unable to scale good pilots into integrated, value-generating AI systems. Through an integration of theoretical insights from the literature with first-hand expert views, this work aims to provide practical advice while expanding our understanding of what works and what doesn't in enabling or disabling sustainable AI uptake.

Interview questions were developed from the conceptual framework established in the first chapter and addressed implementation challenges, organizational readiness, infrastructure and data constraints, and AI maturity (Appendix 1). Such an overarching framework offered some consistency in the interviews but still permitted some room in terms of flexibility based on job, industry, and experience level of the participant.

The study is not intended to produce findings that are statistically generalizable but rather to seek emergent themes and learning in a range of organizational environments. It is a methodology well-suited to a field where success or failure will most likely be a function of fine-grained and context-sensitive variables that cannot be captured by quantitative analysis.

2.2. Conceptual Framework

This research adopts a conceptual framework that is based on three dimensions underpinning organizations' adoption and scaling of artificial intelligence: AI maturity, organizational readiness, and technical enablers (see figure 1). These were derived through literature review and are utilised as the analytical lens in interview design and thematic interpretation.

Organizational readiness involves factors such as leadership commitment, cross-functional collaboration, maturity in terms of change, and cultural openness towards adopting AI. Technical facilitators involve data infrastructure, system integrative capacity, monitoring habits, and AI model life cycles management. AI maturity models give a broader understanding of how AI is strategically and operationally integrated in different organizations.

Rather than keeping those elements separate, the model considers their interconnectedness in a consideration that scaling AI success relies equally on organizational alignment and technical capability. This model ensures congruence is established through the empirical analysis and theoretical model so that the practice knowledge is taken into consideration in a research-based coherent framework.

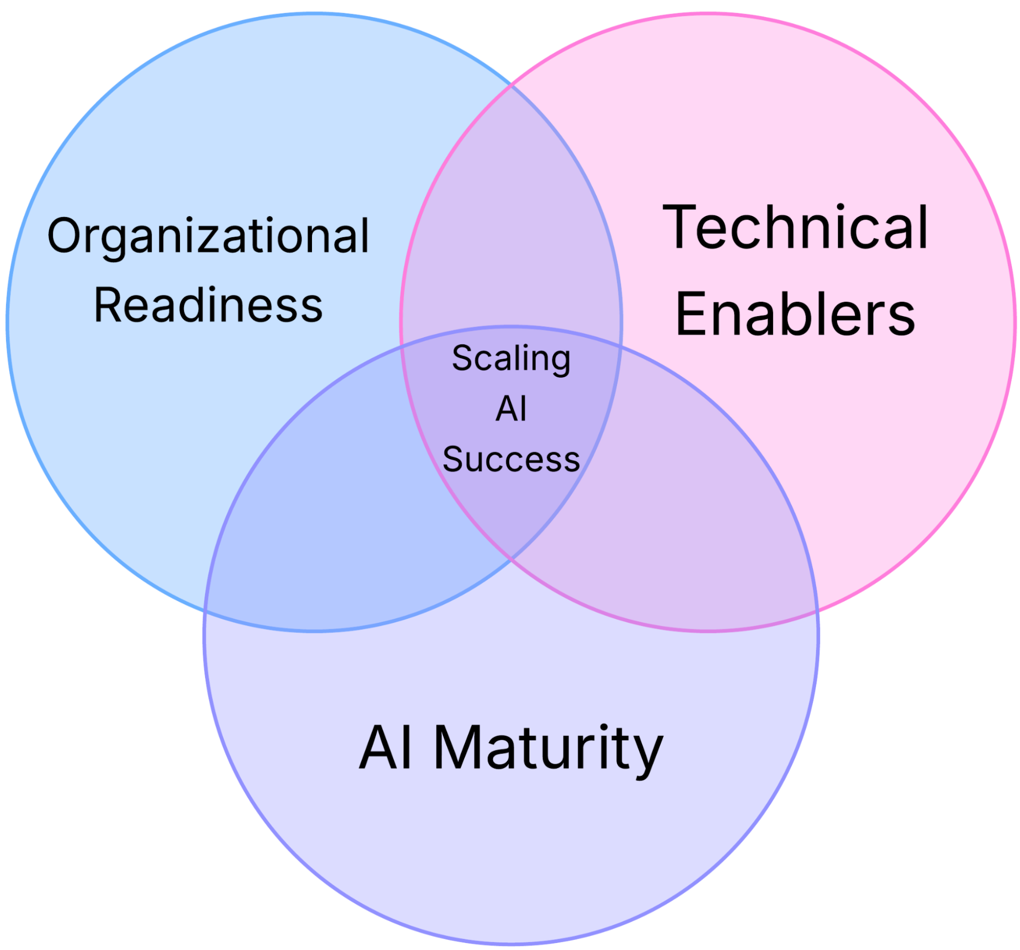

Figure 1. Conceptual Research Framework

Source: Author’s elaboration.

2.3. Data Collection and Interview Process

The study is conducted in a series of semi-structured one-on-one expert interviews selected as most appropriate to gaining in-depth insight into AI implementation organizational and technical realities. It is well-suited to exploratory research, particularly where the aim is to uncover divergent experience-based views based on organizational context.

Interviews are organized remotely, using video conferencing software (e.g. Teams) in order to suit broad nationalities of experts. Interviews are recorded (with permission from the participants) for analysis and transcription. Each participant is informed prior to interview about the purpose of the study, confidentiality of data and voluntary participation.

An interview guide based on the themes emerging from the literature review and conceptual framework has been developed. While the guide offers consistency across interviews, it still allows room for free discussion and flexibility to accommodate issues arising. The key areas of focus in the guide are:

- Strategic alignment and leadership support in AI initiatives;
- Technical and Integration Problems;
- Organizational readiness, competencies and workforce adjustment;
- Factors that enable or impede successful pilots transitioning to production;
- Implications of AI maturity on cross-function collaboration.

Participants are invited to bring cases from their own organizational context, both examples of successful implementation and project examples that encountered problems. In this way, it is possible both to discern common patterns and to capture case-specific, idiosyncratic insights. All of the interviews are anonymized and prepared for analysis.

2.4. Research Population and Sampling Strategy

The population that was targeted in the study consisted of professionals who were directly involved in the application of AI as well as in the organizational context that determines the adoption and expansion of AI technology. The population was restricted to those who held senior or expert roles in which they contributed toward decision-making, implementation, or governing in relation to AI-enabling solutions.

The study adopted a purposeful sampling method, which ensured the selection of participants with relevant experience relevant to the research objectives. The initial participants were sought through the researcher's professional network with a view to gaining access to experts actively engaged in AI-related projects. Snowball sampling was then used as a follow-up method after completing initial interviews, where participants were requested to offer suggestions about other participants who met the selection criteria.

A total of nine semi-structured expert interviews were conducted. This study’s sampling framework included many different sectors: manufacturing, banking and financial services, pharmaceutical and healthcare-related services, beverage manufacturing, management and digital transformation consulting services, information technology services, and energy and utilities. Roles and functions varied from AI leadership and strategy to product and engineering management, process ownership and governance to operational excellence. In order to maintain anonymity and safeguard confidentiality, they have generally been referred to by role and function and not by name and organization. The make-up of this sample is detailed within table 3 (Chapter 3.1).

2.5. Data Analysis Approach

The interview data is analyzed using thematic analysis, a well-established qualitative research technique to identify, code, and interpret patterns in datasets. Thematic analysis allows the researcher to remain close to the data but be in touch with broader theoretical thinking and organizational phenomena.

The analysis follows process adapted from (Braun & Clarke, 2006):

1. Familiarization with data - Transcripts are read and reread to get to know everything about the feedback of all the participants.
2. Coding initially - Descriptive coding applied to portions of text based on relevant issues, experience, or themes.
3. Searching for themes - Codes are then organized into broader themes that demonstrate patterns that are repeated by the participants.

4. Reviewing and refining themes - Themes are refined so that the themes are discrete, coherent in themselves, and assist in analysis.
5. Reporting - Representative quotations and major themes will be used in reporting to demonstrate findings and confirm the interpretation of the results.

Coding will be a blend of inductive and deductive reasoning: initial codes will be guided by the conceptual model and literature review (technical barriers, organizational readiness), but emerging themes are also allowed to be developed from the data themselves. A hybrid approach allows both theoretical coherence and openness to new insights.

2.6. Trustworthiness and Scope of the Study

To ensure research rigor in the qualitative context, this study incorporates the principles of trustworthiness developed by (Lincoln & Guba, 1986). These principles ensure a means of creating quality, consistency, and clarity in the research product and process:

- Credibility is assured through carefully choosing the participants so that everyone interviewed is in today's and current proximity with AI uptake. Use of semi-structured interviewing allows in-depth questioning, with literature triangulation helping in ensuring interpretive accuracy.
- Transferability is facilitated by reporting contextual information about the roles of the participants, industries, and organizational settings. Even if the results are not expected to be generalizable at a statistical level, they are potentially transferable to other organizations struggling with scaling AI efforts.
- Reliability is created by making the research process open and traceable, such as by using an interview guide, coding scheme, and analysis diary. This enables others to follow the steps in the methodology taken and see how the meanings were created.
- Confirmability is established through minimizing researcher bias through careful documentation, direct quotes of the research participants, and constant reflexivity in analysis and coding.

The scope of this research is confined to the organizations that set off AI initiatives with the aim to scale or operationalize them, or move beyond pilots. Only organizational and technical aspects of this transformation are explored in this research, while broader social, ethical, or legislative ones are neglected unless those come up naturally in the field interviews. This research is not trying to measure the success of implementation but explores self-assessed enablers and barriers based on expert views.

3. RESEARCH

3.1. Participants and interview context

In this section, the empirical data used for the purpose of the analysis will be outlined, along with a summary on the interviewees. The results being presented in this work are the culmination of semi-structured expert interviews with professionals working on AI strategy, delivery, management, or operational transformations. The recruitment of the interviewees sought representation amongst those interacting with the business side, as well as in different types of organizations. In order to ensure confidentiality, the participants have been given codes in terms of their roles. Table 3 below is a summary table of the interview participants classified according to their organisational fields/industries.

Table 3. Overview of Interview Participants

#	**Title**	**Company Field**
1	Senior leader responsible for AI and data-driven operational performance	Global Business Services - Manufacturing
2	Process owner responsible for end-to-end operational processes	Banking
3	Product manager responsible for data and AI-enabled solutions	Manufacturing
4	Senior consultant and delivery lead for AI and transformation initiatives	Management & Digital Transformation Consulting
5	Senior finance leader involved in digital and AI-enabled transformation	Pharmaceutical & Healthcare IT
6	Engineering manager responsible for AI system development and deployment	Consulting & IT Services
7	Global process management leader overseeing standardized operations	Beverage Production
8	Operational excellence lead responsible for process optimization initiatives	Pharmaceutical Industry
9	Senior automation leader responsible for intelligent automation initiatives	Energy and Utilities

Sources: Author's elaboration based on interview data.

3.2. Artificial Intelligence Adoption And Use Cases

In this section, the empirical results achieved through the expert interviews carried out will be reported. This includes how the organizations in the studied sample use AI and the type of use cases they focus on. Full codebook is provided in Appendix 2 to ensure transparency and traceability of the analysis.

3.2.1. Motivations and Drivers for Adoption

Findings from the interviews suggest that primarily two closely intertwined motivations exist and drive organizational involvement in AI: (1) immediate increases in productivity offered by easily available generative AI tools, and (2) future improvements in process performance offered by domain-specific automation. The first factor is mapped more closely to ease of implementation and a rapid take-up rate among the organizational members, whereas the second factor is linked more closely to organizational capabilities and possibilities regarding scaled-up development from pilot projects. This corresponds to the distribution of coded content noted in the codebook, as the organizational references revolve around categories that consist of General Productivity and Internal Agents, as well as function-type automation, for example, accounts payable automation and fraud protection.

Throughout the interviews, these drivers emerged as a typology of use cases that encapsulates an increasing level of adoption maturity, as listed in the summary in Figure 2. They tend to begin with general productivity tools, progressing towards domain agents, and in certain cases, experimenting with more innovative or novel applications.

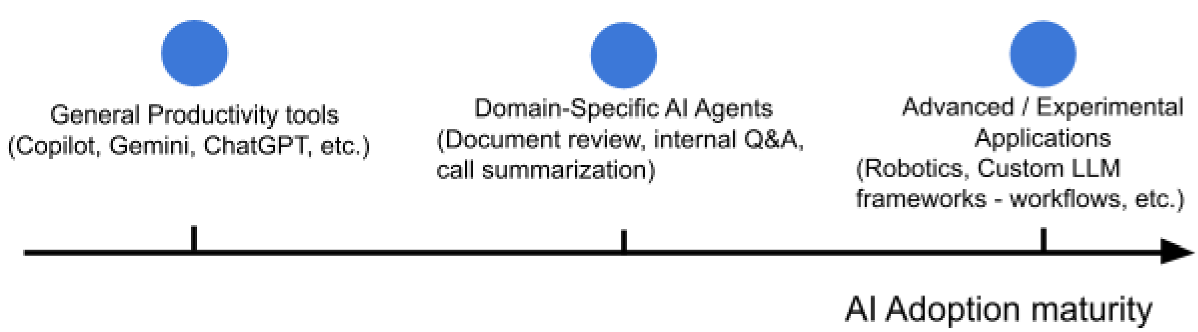

Figure 2. AI use-case categories across levels of adoption maturity.

Source: Author’s elaboration based on interview data.

The typology is used in this section as a guiding framework to discuss how organizations legitimate AI initiatives and the reason why some initiatives stagnate before reaching the level of the productions.

A common justification for the adoption of AI has been related to the expected ability to save the time spent on most mundane tasks while allowing workers to attend to higher priority ones. This mindset was especially observable with regards to applications with an agency-like nature (such as summarization and internal support), which were perceived not as full automation but rather as an enabling action. Coinciding with the first mindset, there was also a second dominant justification related to specific process improvements. The justification for the adoption of AI was made by all participants on the basis of hypotheses about its potential for improved process performance. For instance, one participant explained the process: "…I can do RCA three times shorter… using AI tool," (P8).

At the same time, the need to avoid limiting future adoptability through restrictions imposed by some concurrent trials is emphasized. A very common restriction relates to the challenge of credibly proving business value to justify larger-scale improvements and win acceptance from different business units. Furthermore, it appears from the interviews that mere visibility and novelty do not necessarily facilitate advancement to persistent usage. The comment by one participant that "the robot dog concept didn't go past the test pilot stage" (P1) exemplifies novelty projects getting stuck at the experimental phase if it relates to an ineffective use scenario or failing to prove value at such an early stage.

In conclusion, the evidence shows that the adoption of AI technology is driven, on the one hand, by the benefits of immediate increased productivity and benefits to process performance; but, in order to progress beyond the pilots, it is necessary to convert these anticipated benefits into tangible value.

3.2.2. Stages and Maturity of AI Adoption

The data gathered through interviews suggests that the process of AI adoption typically takes a series of staged steps involving progression from idea conceptualization and experimentation through validation of the idea or concept through proofs of concept, and in a few instances, scaling. It can be seen that the dominant codes are centered around terms such as

"Experimentation importance," "POC importance," "Pilot projects," and perhaps even "Pilot tests."

A clear trend is that there is a desire to prove both the suitability of the use case and the acceptance before committing to scale. A comment from one of the organizations illustrates this progression from considering a potential solution to implementing it after what is essentially "rigorous proof of value proof of concept type pilot … to validate the use case and the solution." (P3). This suggests that the question of maturity is not simply answered by the potential to be implemented or the potential to be useful, but also the potential to prove its usefulness, desireability, or fit within the organization on a controlled basis. In another example regarding the definition of the gate, the gate itself is stated to be the point at which the assessment is made regarding scalability and demand: ".assess the outcome. was this useful. is it something that we can scale. is it something that people want. and sometimes the answer is no" (P3).

The evidence also indicates that scaling becomes contingent on budget ownership or operational sponsorship often as a factor to consider when undergoing the maturity process. As one of the participants indicated, his team was able to move their solutions to the next level using the activities involved in proof of value, and this involved the operational company judging their solution as valuable as they sought to benefit from return on investment as a factor for scaling to take place.

At last, the participants emphasized that pilot activity in its own form comes in different shapes and sizes, ranging from testing with carefully established participant groups to small-scale experiments with the aim of impact demonstration before expansion. For instance, this was demonstrated by one participant who reported carrying out the pilot activity through a controlled pilot with an established participant group: "…a focus study with 50 individuals… as a controlled test" (P1).

In summary, the findings reveal that AI adoption maturity is effectively more of a stepwise process, whereby scaling only takes place if proof of value and pilots are consistent with sponsorship and funding within the organization; otherwise, these initiatives would not progress beyond the experimental/proof of concept and Pilot stages to the enterprise-wide adoption level.

3.3. Organizational Readiness for Implementation and Scaling

The data from the interviewees show that a state of readiness for AI implementation within an organization correlates more with the capabilities of the organization to (a) identify projects for AI implementation, (b) obtain support and investment, (c) adhere to cross-functional alignment during implementation, and (d) prepare employees for changes when compared to the capabilities of a particular technology within the organization. This also applies to the codes illustrating priorities and experiments, organizational capabilities (such as development of capabilities and champions), and organizational implementation challenges (such as lack of capabilities and employee fear).

3.3.1. Prioritisation and Selection of AI Initiatives

One of the constant processes that have been illustrated in the interviews relates to the use case prioritization and ordering process preceding significant investment. This helps in the adoption of AI, rather than actioning it as a homogenous "roll-out" process. This can be exemplified by one of the interviews that cited the use of a scoring process that compares the complexity involved with the gain in efficiency to result in the ordering of implementation "you are creating a table. you have some kind of score how easy it is to set up. how big the efficiency" (P2)..

Simultaneously, selection pressure was posed as a set of pragmatic heuristics prioritizing high-impact opportunities given resource constraints. A research participant encapsulated this as the “80/20 rule,” (P8) stating, “We always pick the biggest ones and weigh potential as the driving factor for advancement” (P8). It appears as if organizational preparedness is partly encoded within decision-making practices safeguarding against effort diffusion over multiple proofs of concept and rallying behind use cases which are simultaneously implementable and impactful.

3.3.2. Sponsorship, Funding, and Organisational Timing Constraints

Among all the interviews combined, readiness is also related to the ability of the organization to move from an idea to an approved business case and its funded implementation within an agreed-upon timeframe. One respondent associated the implementation speed with the speed of technological change when claiming that “a lot of the work is getting the idea, the business case and the right technology to get the funding. .. before the technology is too old" (P8). There is an understanding here that scaling is more about technology, governance, and investment because any delays in organizational implementations may impact the value of implementing technologies and models that are continuously evolving.

Furthermore, the results of the interviews situate pilots as a governance tool bridging the gap between experimentation and commitment. Several interviewees described the concept of proof validation as follows: "rigorous proof of value proof of concept type pilot… to validate the use case and the solution"(P3), but this also comes with the possibility to "sometimes the answer is no" (P3). The readiness pattern is significant in the fact that it supports learning but also filters the projects not creating enough value to be deployed within an enterprise.

3.3.3. Capability Alignment and Realistic Expectations for Performance

Readiness constraints are often coupled with a lack of capabilities and overly optimistic views of AI capabilities. An informant described a common misconception that a firm can "just put AI, give it procedures for your process, and it will do it for you"(P4), and then dismissed this hypothesis: "That's absolutely not happening. Not anytime soon. it's just not nearly as smart as people think it is"(P4). This exemplifies a point that readiness entails interpretive capabilities-that is, an understanding of system limitations on the part of the organization, because overly optimistic expectations can result in improperly specified requirements and disappointment during implementation.

Furthermore, there is an implication from this content that cross-functional coordination is part of readiness, particularly when process knowledge and system knowledge do not match with external knowledge. In one text, lack of collaboration is presented as expensive, such that process owners and system owners and IT people and consultants "are on the same call and no one knows what they're talking about"(P8), which is identified by the informant as a situation to avoid on account of expense. In such a case, readiness covers more than mere technical ability.

3.3.4. Workforce Acceptance and Change Management

The findings suggest that workforce readiness is an important factor in moving past limited experimentation. Resistance and fear were reported as being especially pertinent within rule-based domains such as finance and legal, in which there was seen to be a direct substitution for human labor by the presence of AI. A research participant stated: “Many people are very concerned: this is going to replace their job” (P1). A participant was also reported as saying: “It will steal my work”(P2), in an effort to somehow undermine the application. The research also finds that external assistance is often supplemented by internal preparedness, especially in the absence of previous internal experience. In this respect, “an AI outcome manager”(P3) type of position has emerged on the vendor side, which acknowledges that vendors perceive organizational enablement as part of the service offering. Another informant told me more directly that even with “the smartest people” it is “really hard” without prior experience with adoption, and that “most of them…can really benefit from a good partner” (P4). In this respect, preparedness is seen to incorporate partnership expertise, particularly during early adoption phases of system expansion.

In general, the empirical evidence indicates that readiness in organizations for the implementation of AI involves well-structured prioritization approaches, well-timed governance and funding, realistic views of capabilities, and change management focused on the workforce. The results also indicate that partner ecosystems can mitigate the lack of experience within an organization to a certain extent but do not eliminate the need for decision discipline and adoption legitimacy.

3.4. Technical Capabilities

The data shows that the viability of expanding AI-related projects has been significantly influenced by certain technology capabilities and limitations. Even if the proof of concept for AI projects proves immediate functionality potential, the productionization stage can only have a sound technology foundation involving scalable infrastructure, sound model behaviors, safe data management practices, and sound maintenance processes.

3.4.1. Infrastructure Scalability and Deployment Feasibility

There is a fundamental technical competency that has to do with the ability to provision infrastructure that can scale to support additional usage beyond the constraints of a pilot scale. The subject of infrastructure provisioning and its inability to scale the demands of running additional usages of the infrastructure came up as a factor, and it seems that it is expected that a technical infrastructure would need to scale its provision of generative model usages, as one subject implied, “when like 20 people started using it, it choked, you have to have an LLM in the cloud. And, of course, that has its own problems, but that's a different discussion"(P4). This suggests that technical competency has requirements that go beyond the selection of a model that runs well.

Firstly, in terms of capabilities, this implication of the research finding is that for scaled adoption of AI, what is needed for organizations is (a) the architecture of their infrastructure should not create bottlenecks, and (b) the adoption mechanism should ensure the solution scales for users. The inclusion of the codified reference “IT infrastructure limitations” in the codebook adds evidence to the implication that even if the concept is feasible, infrastructure limitations keep surfacing as barriers.

In sum, it is apparent from these results that the ability to support scaling and flexible deployment for AI is necessary to prevent projects from being limited to pilots.

3.4.2. Data Readiness and Process-data Dependency

The second technical requirement is the readiness of organizational data necessary for implementing the AI use cases. The theme of “Data readiness” came up with 7 citations in 5 interviews, indicating this is a prevalent concern across participants as opposed to a concern specific to a particular context like a business organization. The evidence of concerns about data readiness indicates that technical feasibility can be subject to the availability of the data.

It can be further supported that such a dependency is also tied to the coding category that matches process feasibility and duration in projects to “data availability and accessibility,” which suggests that even when a process can be qualified as being automatable, due to data-related factors, the timeframes associated with implementation and feasibility can still be constrained. It serves a gatekeeping function when data readiness in an organization lacks maturity in data access.

The technical capability of data readiness appears to be an enabling factor whose absence causes either delays or reduced projects due to inadequate access, quality, or availability for effective automation.

3.4.3. Reliability, Hallucinations, and Verification Requirements

In addition to the issues of infrastructure and data, the following concerns regarding the reliability and correctness of model outputs, especially for those use cases that would be operationally or regulatorily intolerant of mistakes, emerged. As evident by the coded data, for one participant, hallucinations and the constraints on adoption that this means for the technology, since “the hallucinations are happening”(P2), complete substitution cannot be accomplished. They further comment that for the audit-driven environment that they operate in, “I can't show. 30% mistakes”(P2).

Likewise, another respondent articulated that the primary question, not just one related to hallucinations, but one related to trust within the output itself, was encapsulated within the following statement: “at the core it’s still like trusting the output”(P3). This indicates that the question of whether the system is technically capable encompasses both the selection of the model itself, as well as the process validation procedures that would produce evidence thereof.

A distinction arising in the data is between off-the-shelf use and more controllable engineered systems. For example, one participant stated mitigation can be done in multiple ways to “make the tool not hallucinate”(P4) but also stated that total control can only be assured through customizing the tool, which adds costs as it “requires infrastructure”(P4). The implications are apparent in the finding that to mitigate in more risky domains, there will be a need for architectures heavy in engineering rather than off-the-shelf use.

Results indicate reliability concerns-primarily hallucinations-correlate to levels of verification requirements and trust; scaling is more plausible when a certain level of testing and engineering competence exists to add to controllability and minimize error exposures.

3.4.4. Security Constraints and Cloud/on-premise differences

The questions of security and confidentiality were also issues in a number of interviews and related to 3 direct references within 3 different interviews under the category of “Security concerns.” The issue of security has been a fundamental criterion of evaluation for AI technology, particularly in connection with confidential information related to a company and/or human resources issues. This was reflected when one of the interviewees indicated that clients are “focused very much on. how the data is secured”(P6), and another was interested in how confidential HR information would be secured, and another directly referenced on-premise vs. cloud technology and how tools are “usually. cloud-based”(P9).

Taken together, this evidence indicates that technical competence must include the ability to provide a secure operating environment and risk management framework that satisfies the risk tolerance of the organization. Furthermore, the necessity for cloud-based tool usage as a technical competency creates a clear conflict, in that the requirements of the operating environment may constrict what data and business workflows are enabled for processing in the cloud.

3.4.5. Maintenance and Rapid Technology Change

Interviews show the importance of technical preparedness being a dynamic rather than a one-off process by describing it as follows: “Monitoring and then maintaining a system over a lifetime is a requirement. It’s a requirement. Because of course, there’s going to be continuous monitoring, as opposed to deploying a solution and letting it go”(P3).

Likewise, another individual described “data drift”(P1) by identifying how technical maintenance can be the same as other software solutions since it needs “constant maintenance”(P1).

In addition to drift and monitoring, data also reflects instability in the technical system. A respondent spoke to how certain changes to platforms have simply undone months of accumulated configurations and indicated that “all my work is gone. I have to start. from scratch”(P6) following a vendor shift. This suggests that technical quality is more than simply maintaining one’s own toolset and adapting to changes at the vendor/platform level.

The evidence indicates that the scaling of AI needs an aspect of the lifecycle phase related to observing, sustaining, and adapting to rapid changes in technical environments, otherwise known as diminishing solutions.

3.5. Artificial Intelligence Adoption Maturity Levels

Previous conceptual frameworks view the evolution of AI Maturity as a move away from ad-hoc experimentation and towards more managed and enterprise-wide approaches. The empirical results are broadly in line with such conceptual frameworks; yet, it would seem that such maturity cannot be aptly conceptualized as merely a linear construct. This emerges because:

(a) it varies unevenly in its dimensions, and (b) it tends towards self-overestimation insofar as experimentation visible to the public eye can be taken as a surrogate for institutional preparedness.

3.5.1. Overall Positioning in Maturity Models

Individually, the organizations are generally positioned closest to the "pathseeker" stage of the process, compared to the "transformer" or "leader" stage on the Deloitte's maturity model. This is because there are signs of internal pilots, departamentalization, as well as the beginning of standardization (solution reuse), without much signs of the full integration at the more holistic level.

In all instances, participants discussed the use of general-purpose productivity tools and more exploratory endeavors like internal assistants or agent-based systems. These tendencies suggest that AI is now viewed as having strategic significance and that an initial awareness phase or isolated proof-of-concept is ending. At the same time, the more integrated pattern of organizational use typical of higher levels of maturity-established processes, steady usage across units, and defined scaling practices-is less common.

The group of test organizations tends to show a good amount of "beyond-awareness" maturity in relation to pilots and early standardization, yet no institutionalized forms of later-stage maturity.

3.5.2. Asymmetry Across Maturity Dimensions

One of the most salient patterns that emerges is that maturity tends to progress asymmetrically along these different dimensions, with technology and data capabilities regularly growing more advanced in terms of AI than governance structures, talent factors, and cultures. In this respect, technological maturity is represented by the adoption of more sophisticated methods, such as large language models, with the aim of developing a reusable platform or internal models to accelerate the next project. At the same time, more mature instances tend to aim to incorporate AI with internal knowledge bases and business data despite technological constraints.

In contrast to this rapid process of change and improvement, governance and employee capabilities are portrayed as undergoing slower development. Governance is viewed as an evolutionary process whereby it exists by intention and to some degree through practice but remains less formally established through formalization and monitoring. In much the same way, leadership involvement and subject passion were present but sometimes lacked employee capabilities and understanding.

This multi-dimensional approach corresponds to the operationalization of maturity for at least one of the participants, which defined organizational routes of execution as being dependent on ‘maturity readiness’ and ‘competency level within the system and process,’ thereby indicating that maturity was not a theoretical construct but had a role to play in execution.

The results reveal imbalance in maturity levels on different dimensions, with relatively faster development in technology, compared to governance, skills, and culture dimensions. This directly impacts implementation routes.

3.5.3. Differences in Challenges Across Maturity Levels

The nature of barriers changes with maturity, according to what the interviews have brought to light. For lower maturity, the kind of barriers will be organizational: lack of clarity about where to begin, a lack of certainty about what ‘good’ AI looks like, difficulties regarding transitions from prototypes into formalized, evaluated projects. Here, the project will typically stay pilot-oriented, informal, until more complex technical issues such as architecture design/data will become relevant as a project is poised to start scaling.

In higher maturity environments, there were more structured processes to suggest, select, and endorse AI endeavors, such as processes related to portfolio distributions in an agile way (like internal “bidding”(P8)). In these situations, the underlying question shifts from whether to use AI to ways to apply it in alignment with strategy, feasibility, and value. However, rather than reducing obstacles, maturity involves moving on to more complex issues, such as legacy system integrations, data governance issues, accessibility, and dependence on limited experts for comprehensive initiatives.

As the level of maturity rises, the barriers to uptake appear to increasingly move from the uncertainty and friction of adopting a new technology to a complex technical scaling problem and from organizational uncertainty to a portfolio problem. They do not completely disappear.

3.6. Project Successes and Scaling Decisions

It appears that the notion of success in the context of AI endeavors uncovers its basis for definition in terms of: (a) the achievement of quantifiably enhanced process performance, (b) organizational buy-in for continued support, as well as financial commitment toward continued application, and (c) risk-controlled deliverability of dependability commensurate with the process context in which it functions. Instead of emphasizing success as an objective property, it was conceived as being an element within the logical basis for an executive-type decision: irrespective of whether an activity can present an irrefutable justification for continued support, such as compelling evidence for continued support.

3.6.1. How Success is Defined

Through the interviews, success was often related to hypotheses about operational performance that could be proven in pilots. For example, success was defined in terms of cycle time reductions, productivity, quality, and accuracy levels often expressly stated as conditions under which managers would approve. For instance, one respondent correlated value realization to productivity improvement based on the controlled pilot, reasoning about approximate productivity savings of “a 2-hour-a-week productivity gain”(P1) among office workers, and concluded the cost of licensing and training would “pay for itself 10 times over”(P1) even if it is not directly apparent until organizational capacities shift.

A second thrust was the idea that “success” is contingent on the relevance of the measures to the function. In the context of sales effectiveness, a respondent described customer use of precisely calculated performance metrics such as cost per visit and the associated return, with continued use dependent on measures of benefit, stating: "Customers are using performance measures like 'cost per visit' and its corresponding return. If we do not see a benefit, then clearly we do not see the reason to pay “x” per user”(P5).

The success of a project for the most part is established as decision-grade value-in other words, as the measurable values of improvements in terms of timeliness, quality, accuracy, and return.

3.6.2. Evaluation Practises and the Use of KPIs

They depicted notions of evaluation that corresponded with the logic of comparative testing, such as quasi-experimental thinking. In regard to customer decisions, for instance, an informant talked about measuring performance in terms that corresponded with controlled comparisons, with adoption contingent upon whether it altered outcomes such as cost or return per visit and whether value added was sufficient to warrant further payment.

In addition, the data shows that thresholds for accuracy and reliability are important success factors, and even more so for processes that require compliance with regulatory requirements. In discussing what constituted successful results for the company processes that took longer to execute, one respondent said that what they required for most processes “95-96 is fine,” but for others it required “almost 100”(P4). Yet another respondent associated the “success rate” with the amount of work that human validators had to perform when results had to be validated for inaccuracies and required correction for the mistakes they entail. "Accuracy is the most important for business"(P9) with greater stress on turnaround time and time savings.

In addition to static evaluation, it can be observed from the coding that some of the participants understand ‘success’ in a more continuous and operational sense, as opposed to a single evaluation point in time, since one of the participants indicated that AI deployment is a process of ‘continuous monitoring’ and ‘not a deploy and let it go’ process.

Success for evaluation functions tends to involve (a) measurable changes to key performance indicators, (b) satisfactory levels of accuracy for the process risk level of evaluation functions, and (c) continuous evaluation function practices to maintain trust after a successful function implementation.

3.6.3. Failure Patterns and why Pilots Stagnate

Rather, failure was not characterized as "technology did not work," but rather as a mismatch between the initiative and the organizational prerequisites for scaled adoption. Three types of failure patterns emerged.

To start with, interviewees warned that poor use case development choices may entrap organizations into prolonged pilots without creating value. “If you choose a bad use case then you’re screwed”(P4), said one participant, introducing failure as an inevitable outcome resulting from earlier choices. Furthermore, failure at an earlier stage may result in push-back from organizations: “if you screw up your first two or three projects. You're going to get a ban. for the next two years”(P4), suggesting that failure to create value at earlier projects may affect legitimacy and resultant investment appetite.

Second, pilots may fail if the outcome indicates either “common sense” as opposed to a practical benefit or when dominated by external restrictions. The problem described a model of a need forecast which worked well, producing a clear result: “common sense would have told you that anyway”(P1), then went nowhere-and became a white paper’-as the organizational limitations (such as inventory, logistics) constrained the use of such a result. It represents one way of failing, which is if the tool can produce useful information, yet the organization does not use this information towards making a value-creating intervention.

The third category of failure is possible by means of non-adoption on a social and organizational level even when the technology is possible from a feasibility point of view. Resistance with respect to the workforce was directly attributed to concerns over job replacement; for example, a research participant mentioned rule-based functions in which workers are “very worried that AI is ‘going to replace my job’’(P1) and gave the example of accounts payable. Another participant mentioned this type of resistance as a risk of active sabotage: “it will steal my work so I’ll do anything to prove that it’s total crap”(P2). This type of research participant evidence would suggest that failure from pilot to scale could be due to factors of legimitacy and adoption rather than technological inadequacy.

Failure can often be caused by: (a) poor choice of use cases, (b) ineffectiveness at turning results into meaningful value under practical conditions, and (c) resistance of staff and stakeholders against routinizing based on pilot feasibility.

3.6.4. Scaling Decisions, what Determines Continuation or Dismissal

The process of scaling takes a gate-keeping decision-making process with the involvement of the pilots as evidence-producers. The process of staged validation has been described by the participants as one whose aim is to “validate the use case and the solution”(P3), with the addition that “sometimes the answer is no”(P3), signifying that the process of non-scaling can actually be expected and rightful. Some of the participants have reaffirmed that there must be a readiness for the potential of disaster when it comes to the process of scaling because one of the informants wrote that model-updates and disruptions can result in the ceasing to function of the tools used.

Scaling should be viewed as a gated process where pilots produce evidence for or against the satisfaction of the criteria (value, accuracy, speed, sustainability, continuity). Projects will scale or be kept as pilots or be re-designed or terminated depending on the achievement of the thresholds.

3.7. Summarizing Empirical Findings

The empirical evidence shows that the process of transitioning from piloting AI to production is characterized by an interdependent process of organizational and technical factors rather than an individual "best practice." The process of adopting AI is primarily driven by accessible and productivity-focused applications during experimentation, while the process to more embedded solutions is linked to the development of decision-grade evidence about value and acceptability and the ability to operationalize the solution.

3.7.1. Core Mechanism of Enablers and Barriers

Throughout the dataset, a constant challenge in terms of scalability being a lack, not a lack of promising concepts per se, but rather the inability to produce results credible enough toward gaining the commitment of an organization. As regards a successful pilot being discussed in relation to being tested either in relation to workflow/business performance rather than in relation to "technology novelty" in accordance with the notion that value must be shown in functional terms "KPIs… are not based on the technology… they're based on the workflow"(P3), there also are a set of interviews that reveal that even when results are deemed promising, scalability could still be thwarted in cases in which benefits lie in different units, yet expenditures in single-function departments.

The organizational requirements for scaling have a complementary counterpart in the second mechanism: operational feasibility. The technology and management readiness that scaling demands simply doesn't always exist at the pilot stage. This appears in the form of technical requirements like data readiness and instability issues like uncoordinated changes to the production data model, and compliance and regulation issues like GDPR and PII. As a result, successful results in the pilot phase have to be understood as necessary but insufficient: the production deployment now depends on whether the organization has the necessary conditions.

To summarize these patterns, Table 4, combines the cross-cutting themes by identifying the top barriers and facilitators with the six recurring themes. These themes are connected to the organizational capabilities that should be addressed to improve the chances of scaling success. This allows the data to be summarized and presented clearly on (a) the factors that are often hindering and facilitating scaling, and (b) the implications that address the research objective of what the organizations need to improve to have a successful implementation of AI.

Table 4. Barrier / Enabler matrix combined with Capabilities to Strengthen

Dimension	Barrier	Enabler	Capabilities to Strengthen
Use-case legitimacy	Poor use-case selection. Unclear what good AI use looks like, slow approval.	Disciplined use-case selection, push from champions, clear business case formation.	Better processes for agile project selection, sponsorship and funding routines.
Value proof	Not definitive pilot results, KPIs not tied in to the workflow	Clear KPIs and good understanding of the desired result.	Better KPI design, evidence standards, value tracking. Cross-unit alignment on benefits.
Risk and Compliance	GDPR/PII constraints, unclear data ownership/access.	Early compliance/risk involvement. Well governed data model.	AI/Data governance, privacy/risk routines.
Technical Readiness	Infrastructure unable to handle operations “chokes”. High error/hallucination rates.	Scalable infrastructure (oftentimes cloud-based). Good Data foundations.	Lifecycle engineering (testing, monitoring), strong data teams and governance.
Adoption and routinisation	Workforce anxiety, misunderstanding of AI capabilities, lack of talent.	Trainings and workshops. Leadership push, change management.	Change management capability, AI literacy and expectation management.
Sustainability, Upkeep	Deploy and forget mindset, no clear owner after deployment, model drift.	Continuous monitoring, defined ownership, maintenance practice.	AI operating model including ownership, maintenance, improvement.

Sources: Author’s elaboration based on interview data.

In the next section, these dimensions are converted into a stage gate model: they will be repurposed as “go/no-go” decision gates where AI projects will be either piloted, redesigned, or moved on into production.

3.7.2. Decision Gates for Scaling AI Projects

When the results from each section are combined, a stage gate approach where pilots play a role as a mid-stage gate for the productionization process is indicated. In such a stage gate approach, a project gets to the stage of production not based on the success of a pilot alone but based on the continued evaluation criteria as to its value for money, risk acceptability, technical viability, uptake, and sustainability (Figure 3):

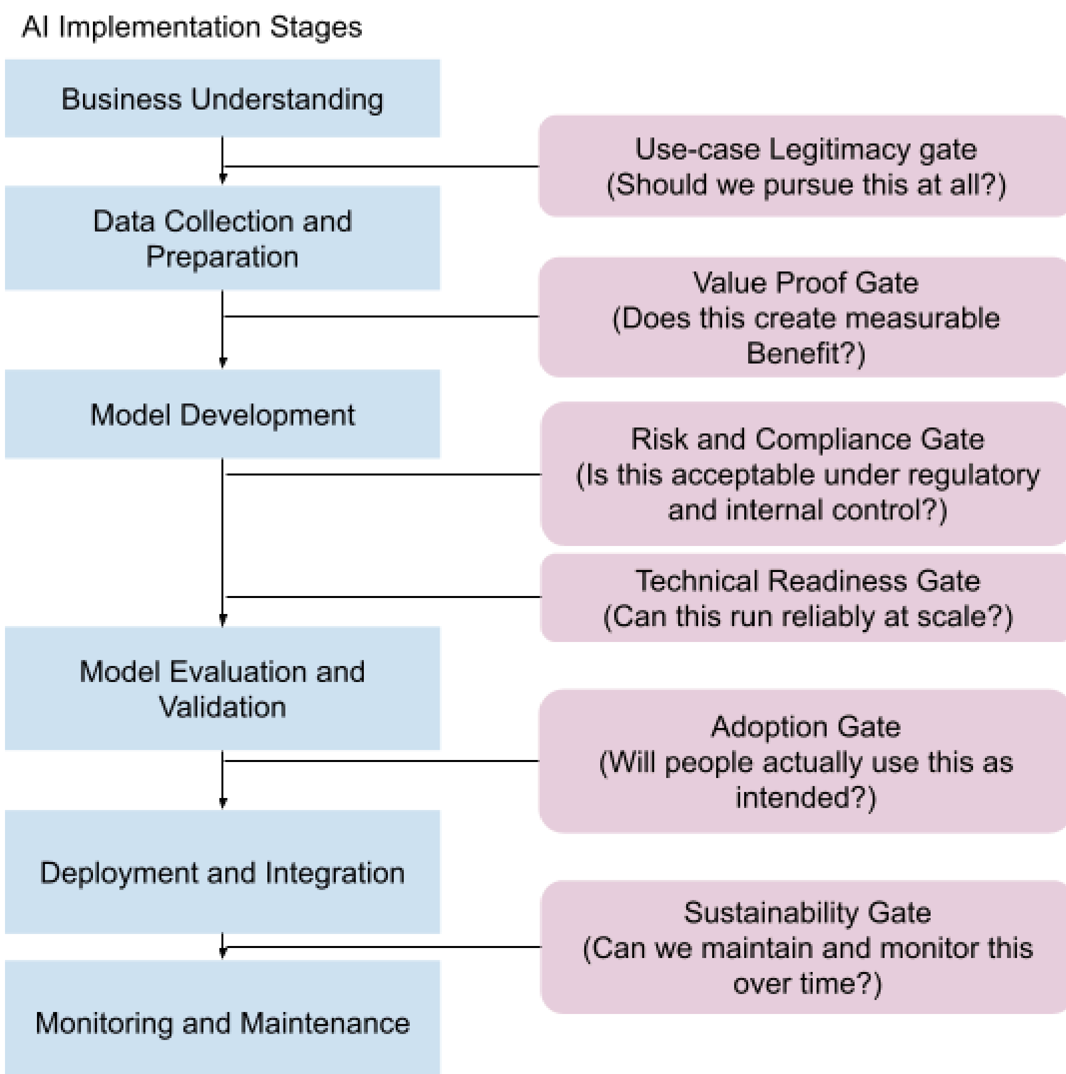

Figure 3. Decision gates across AI implementation stages

Source: Author’s elaboration based on literature review and analysis of expert interviews

Though Figure 3 above links gates to conventional implementation phases, the findings above show the actual implementation practiced in AI initiatives is often done iteratively (e.g., agile implementation), suggesting gates can be dealt with simultaneously or in a different order.

However, scaling the initiative requires meeting all the conditions at the gates.

1. Use-case legitimacy gate (Should we pursue this at all?): Lack of early success can be a product of substandard use-case development. A participant reflected that if a company "picks a bad use case then you're screwed"(P4), and explained that failed early stages can come back to haunt them in terms of maintaining validity for several years.
2. Value proof gate (Does this create measurable benefit?): Participants explained the decision to continue in terms of willingness to pay and tangible return (e.g., "cost per visit and what is the return… we don't see any benefit so we don't want to pay…"(P5)). In the more confident domains, the metric of value included tangible increases in productivity (e.g., "a 2 hour a week productivity gain… pays for itself 10 times over"(P1)).
3. Risk and compliance gate (Is this acceptable under regulatory and internal control expectations?): Scaling is difficult when risk exposure is high or in cases where associated routines are not clearly understood; participants mentioned scaling difficulties explicitly in the context of GDPR/PII or organizational rights. This becomes crucial when pilot usefulness comes into the picture.
4. Technical readiness gate (Can this run reliably at scale?): Technical restrictions are more stringent during production efforts. Some of the obstacles mentioned were infrastructure restrictions ("when 20 people started using it, it choked. you have to have an LLM in the cloud"(P4)) and reliability concerns like hallucinations ("hallucinations are happening. I can't show. 30% mistakes."(P2)).
5. Adoption and routinisation gate (Will people actually use this as intended?): A fully functional system can still be absent due to non-adoption in the social setting. This was observed with resistances against systems because people fear losing their jobs, with some resistances aiming to discredit technologies they feel will harm their work ("it will steal my work so I will do anything to prove that it is total crap"(P2)).
6. Sustainability gate (Can we maintain and monitor this over time?): The success of production lies within the continual state of operations as opposed to the process of

implementation. Such gate-driven synthesis helps to understand why there can be many pilots within an organization but not enough to attain the scale seen in production.

Pilots can succeed in the “usefulness” gate but fail in the gates related to proof of value, compliance, readiness for technology, adoption, or sustainability.

3.7.3. Artificial Intelligence Adoption Maturity as a Moderator for Implementation

The findings also suggest the presence of moderation by the “maturity” dimension in the set of barriers. When the AI adoption maturity is lower, it is likely to face barriers in terms of uncertainty in the organization, poor governance, and the ability to form decisions. In contrast, in a mature context, the barriers will emerge in terms of architecture, complexity of integration, control of risks, and maintenance of the lifecycle. This corresponds with the empirical findings, suggesting overall ‘maturity’ to be a multi-dimensional construct dependent on “competency level in the system & the process”(P8).

In practice, maturity-related reasons account for why similar results from pilots yield different outcomes for an organization: more mature governance and capability structures mean higher chances for pilots to inform resource allocation and plans for implementation, and conversely.

In short, the experimental evidence offered by this research tends to indicate that crossing a series of decision gates, including the formation of a legitimate use case, providing value, living up to expectations for compliance and risk, accomplishing readiness and reliability, forging adoption with the workforce, and remaining effective with monitoring and maintenances, is a factor of crossing a stage of organizational maturity issues, which tend more towards adoption issues rather than issues of infrastructure and reliability. This synthesis serves as a foundation for the following conclusions and recommendations.

3.7.4. Theoretical Contributions

This thesis contributes to organisational AI implementation research by developing an improved conceptualisation and explanation for pilot to production. On the basis of previous research identifying AI implementation to be dependent on the conjunction of technological, organisational, and environmental factors (Chatterjee et al., 2021; Weber et al., 2022). This study builds on scaling with multiple implementation gates to achieve organisational alignment before progressing past pilots.

First, the results enable the operationalization of the pilot-to-production gap through the identification of lifecycle needs that most often negatively impact scaling: data quality/integration issues, lack of explainability, and, significantly, a lack of monitoring and retraining once deployed. In this way, this study can be seen to expand existing literature on production monitoring/lifecycle management studies (Klaise et al., 2020) to explainability needs for trust adoption (Arrieta et al., 2019).

Secondly, study brings nuance to adoption theory by revealing that AI maturity does have moderation effects, where the relevance of barriers changes from basic readiness factors to scale and optimization factors once the organization has accumulated structure, processes, and prior experience. This improves prevailing models of adoption by explaining when given barriers and enablers have been most salient (Chatterjee et al., 2021).

Third, it shows how transparency, accountability, and regulatory risk are incorporated within the reasons for scaling, indicating that privacy and governance factors are integral, as opposed to being marginal, towards creating technical value within production (Cheong, 2024). Overall, the study links capability explanations for implementation within lifecycle engineering demands, providing a more specific theoretical framework for understanding why there are so many AI projects that are piloted but never scaled (Weber et al., 2022).

CONCLUSIONS AND RECOMMENDATIONS

1. It is argued in this study that the scaling of AI post-proofs-of-concept/pilots is limited in the main by multi-criteria decision-making rather than the proof-of-concept itself. The empirical findings for this study's hypothesis indicate that projects alone are insufficient for productionization; the results of scaling are determined by the continuous evaluation of-legitimacy, value, risk/compliance, technicality, adoption, or sustainability of use-cases.
2. A stage gate logic is unearthed that is complementary to typical models for the implementation of AI. Despite the typical linear model for the implementation stages being data preparation, development, evaluation, deployment, monitoring, starting with business understanding, data generated through interviews show that organizations' execution of scaling is instead a progression through a set of 'go/no-go' gates that span those stages.
3. Being legitimate is a precursor for success, acting as an upstream element that promotes success downstream. It has been evident that poor use-case validation can result in poor long-term results, reducing confidence in AI initiatives when failure happens in the initial stages. It indicates that AI initiatives act as a filter when their initial stage is involved.
4. Value proof is the key economic gate that can be interpreted in practice as WTP in principle-through value logic for benefits such as productivity or cycle-time reduction. Projects remain stuck if value cannot be demonstrated convincingly enough to get funding support and collective agreement across units.
5. Risk and compliance constraints are definitely not secondary factors because they define a scale ceiling. The results have shown that concerns relating to the protection of personal data under the GDPR/PII regulations, restrictions on ownership and access, and uncertainty about the responsibilities for controlling the data set a ceiling on scale even when pilots are deemed useful.
6. Technical viability influences it most in the transition phase, when it moves from use to wider adoption. It has been observed that scalability imposes tougher requirements in terms of infrastructure, reliability, and performance requirements such as capacity

metrics, hallucination in models, and production-quality stability when it moves from use to adoption.

7. However, adoption itself is a clear gateway that may pose a bottleneck in scaling, even if the solution is functionally effective. It is clear that, according to empirical data, opposition may arise from concern regarding job replacement as well as deliberate discrediting. In other words, "successful implementation" is a socio-technical process, as opposed to a direct functional process.
8. The stage of sustainability aka maintenance and monitoring will be viewed as the final gate, since it can be asserted that the ownership factor becomes definitive after the deployment stage. Results have shown that the successful execution of productions requires adaptability in maintaining monitoring and adapting AI.
9. Project maturity serves as a moderating factor that influences the relative weights of barriers and enablers. From this case study, there is evidence that barrier and enabler factors differ in their applicability based on the stage of an initiative. Early-stage projects and late-stage projects prioritize experimenting, local productivity improvements, and exploratory adoption as opposed to focusing on governance, reliability, and operating models. Therefore, barriers and enablers should not be treated as static, their relevance is is ongoing depending on maturity and stage of the project.
10. In general, it is found that the empirical contribution of this study is an integrated explanation of scaling failure, according to which pilots may meet usefulness related criteria, but fail at value proof, compliance acceptance, technical readiness, or adoption/routinization/sustainability. The contribution of this study is to translate general criteria of scaling failure into an empirically defined decision tree explaining discontinuation.

Recommendations:

1. Institutionalize a stage gate process for moving from pilot scale to production scale. Develop clear go/no-go gates based on the key gates that have been determined in this research and scale up based on objective evidence.
2. Encourage upfront and disciplined selection and validation of use cases. Ask for a clear statement of the problem, scope of the process to be addressed, assignment of an owner,

and establishment of a minimum value that can be measured before the pilot is approved and before considering scaling.

3. Embark on risk, compliance, and data governance right from the ideation stage. When it comes to use cases with sensitive data inputs or outputs that impact business decisions, early assessments of data access, privacy laws, regulations, and acceptable residual risk must be accomplished to ensure successful pilots are not derailed at the production stage.
4. Make the technical readiness more scalable than just proof-of-concept readiness. Before moving past the pilot phase, it is essential to check the integrations required, the mode of deployments (cloud or on-prem), to ensure that technical debt does not become a hurdle to scalability.
5. Plan adoption and ownership as a part of production readiness. Assign ownership after deployment. Outline the processes for verification and fallback, and assign user enablement accordingly in order to make the solutions part of the operations.
6. Institutionalize sustainability processes post-deployment. Create an operational model of your AI system with a structure of ownership, monitoring of quality, drift, and hallucinations, definition of processes for handling incidents, triggers for retraining, including ensuring systems in production remain in good health post-deployment.

SUMMARY

Vilius Aleksynas

Artificial Intelligence Implementation in Organizations: Barriers and enablers for Transition from Pilot to Production

Final Master Thesis

Academic Supervisor:

Vilnius University, Faculty of Economics and Business Administration

Strategic Management of Information Systems

Vilnius, 2026

Size: 90 pages, 3 figures, 4 tables, 2 appendixes

This is the final master's thesis examining why artificial intelligence (AI) projects tend to be perpetually stuck in the pilot phase and why these projects are rarely scaled from a pilot to a production environment. The research object is the process of implementing artificial intelligence in an organization. This study will provide an overview of the main factors that both hinder and help in the scaling of artificial intelligence from a pilot to a production phase.

A qualitative approach is the foundation of this research. Semi-structured expert interviews have been conducted with nine experts actively involved in the implementation of Artificial Intelligence strategy and change. The results of the expert interviews have been analyzed via thematic coding, incorporating deductive and inductive themes, starting from the literature and including findings starting from practical experience.

The main hard result for the thesis is an integrated story line for the pilot-to-production divide as a decision process at multiple gates and not purely technical for when to deploy. Empirical evidence indicates that scaling usually involves sequential 'go/no-go' gates for case legitimacy, value, acceptance and compliance, technical readiness and reliability, adoption and routinization, and finally sustainability through monitoring and ownership.

SANTRAUKA

Šiame magistriniame darbe tyrinėjame kodėl dirbtinio intelekto projektai (DI) užstringa savo pilotinėse stadijose. Tyrimo objektas yra DI projektų diegimas organizacijose ir jų plėtimas iš pilot projektų. Šiame darbe pasimatys pagrindinės priežastys kodėl projektai stoja pilotinėse stadijose, koki veiksniai tai nulemia, ir kokie veiksniai padeda plėtoti šiuos projektus.

Šiame darbe naudojamas kokybinis tyrimas atliekant pusiau struktūruotą ekspertų interviu. Buvo atlikti 9 interviu su įvairių sričių ekspertais, kurie visi turi patirties diegiant ir taikant DI projektus. Interviu buvo transkribuojami ir koduojami, iš kodų sudaromos temos kurios buvo pradėtos iš literatūros apžvalgos ir papildytos iš atlikto tyrimo.

Pagrindinis tyrimo rezultatas šiame darbe yra gaunami DI projektų kūrimo vartai, kurie kartu apima ir techninius ir organizacinius iššūkius. Tyrimas parodė, kad dažnai vystant projektus reikia pereiti šiuos vartus gaunant teigiamą arba neigiamą atsakymą ir įrodant projektų svarbą bei legitimumą, kuriamą vertę, atitikimą reguliavimams, techninius reikalavimus, patikimumą, realizaciją, tvarumą, jų priežiūrą ir nuosavybę.

REFERENCES

ABM. (2023, June 5). Role of AI and ML in Programmatic Advertising | AndBeyond.Media. *AndBeyond.Media Blogs*. https://www.andbeyond.media/blogs/role-of-ai-and-ml-in-advertising/

Ajuzieogu. (2024). *AI Data Quality and Bias: Challenges, Implications, and Solutions in Modern Machine Learning*. ResearchGate. https://doi.org/10.13140/RG.2.2.25830.02880

Alaa, M. (2024). *AI in e-Commerce: Smart Recommendation Systems to Boost Revenue | ProfileTree*. https://profiletree.com/ai-in-e-commerce/

Aldridge, A. (2023). *BofA's Erica Surpasses 1.5 Billion Client Interactions, Totaling More Than 10 Million Hours of Conversations*. Bank of America. https://newsroom.bankofamerica.com/content/newsroom/press-releases/2023/07/bofa-s-erica-surpasses-1-5-billion-client-interactions--totaling.html

Almeida, D., Shmarko, K., & Lomas, E. (2022). The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: A comparative analysis of US, EU, and UK regulatory frameworks. *AI and Ethics*, *2*(3), 377–387. https://doi.org/10.1007/s43681-021-00077-w

Arrieta, A. B., Díaz-Rodríguez, N., Ser, J. D., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2019). *Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI* (No. arXiv:1910.10045). arXiv. https://doi.org/10.48550/arXiv.1910.10045

Atom. (2024, January 8). *AI in Logistics: Smart Route Planning for Efficiency*. Ever Efficient AI. https://everefficient.ai/posts/ai-in-logistics-smart-route-planning-for-efficiency/

Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M., & Smith, D. M. (2023). Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Internal Medicine*, *183*(6), 589–596. https://doi.org/10.1001/jamainternmed.2023.1838

Bathaee, Y. (2018). *The Artificial Intelligence Black Box and the Failure of Intent and Causation*.

Bernal, J., & Mazo, C. (2022). Transparency of Artificial Intelligence in Healthcare: Insights from Professionals in Computing and Healthcare Worldwide. *Applied Sciences*, *12*(20), Article 20. https://doi.org/10.3390/app122010228

Bhartiya, S. (2024, March 5). *Artificial intelligence multidimensional impact applications across diverse fields*. SciSpace - Paper. https://doi.org/10.58532/v3biai4p3ch3

Bojinov, I. (2023). Keep Your AI Projects on Track. *Harvard Business Review*. https://hbr.org/2023/11/keep-your-ai-projects-on-track?

Borczuk, A. (2024). *Meet COIN: JPMorgan's Efficiency Wizard*. Product Monk. https://www.productmonk.io/p/meet-coin-jpmorgan-s-efficiency-wizard

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies* (First edition.). Oxford University Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cave, S., & Dihal, K. (2018). Ancient dreams of intelligent machines: 3,000 years of robots. *Nature*, *559*(7715), 473–475. https://doi.org/10.1038/d41586-018-05773-y

Challapally, A., Pease, C., Raskar, R., & Chari, P. (2025). *STATE OF AI IN BUSINESS 2025*.

https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf

Chalutz Ben-Gal, H. (2023). Artificial intelligence (AI) acceptance in primary care during the coronavirus pandemic: What is the role of patients' gender, age and health awareness? A two-phase pilot study. *Frontiers in Public Health*, *10*. https://doi.org/10.3389/fpubh.2022.931225

Chatterjee, S., Rana, N. P., Dwivedi, Y. K., & Baabdullah, A. M. (2021). Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model. *Technological Forecasting and Social Change*, *170*, 120880. https://doi.org/10.1016/j.techfore.2021.120880

Cheong, B. C. (2024). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, *6*. https://doi.org/10.3389/fhumd.2024.1421273

Chow, R. (2021). *Dartmouth Summer Research Project: The Birth of Artificial Intelligence*. History of Data Science. https://www.historyofdatascience.com/dartmouth-summer-research-project-the-birth-of-artificial-intelligence/

Crowley, J. (2025). *Banks wrestle to unlock AI value amid 'data scientist talent drought.'* https://www.thebanker.com/content/067aeb45-28f4-4f0a-820a-e12f1281c4b9

Deloitte. (2022). *State of AI in the Enterprise 2022*. Deloitte United States. https://www2.deloitte.com/us/en/pages/consulting/articles/state-of-ai-2022.html

Deloitte. (2024). *Navigating the tech talent shortage*. Deloitte Insights. https://www2.deloitte.com/us/en/insights/topics/talent/overcoming-the-tech-talent-shortage-amid-transformation.html

European Commission. (2019). *A definition of Artificial Intelligence: Main capabilities and scientific disciplines | Shaping Europe's digital future*. https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

Faheem, M. A. (2021). *AI-Driven Risk Assessment Models: Revolutionizing Credit Scoring and Default Prediction*. *5*(3).

Galkin, W. (2024, April 8). Legal Accountability in AI Failures and Malfunctions. *Galkin Law*. https://galkinlaw.com/legal-accountability-in-ai-failures/

Gallagher, K. (2024, June 17). *Study demonstrates how AI can develop more personalised cancer treatment strategies | University of Oxford*. https://www.ox.ac.uk/news/2024-06-17-study-demonstrates-how-ai-can-develop-more-personalised-cancer-treatment-strategies

GDPR advisor. (2024, January 10). GDPR Compliance in the Age of Artificial Intelligence: Challenges and Solutions. *GDPR Advisor*. https://www.gdpr-advisor.com/gdpr-compliance-in-the-age-of-artificial-intelligence-challenges-and-solutions/

Gençten, N. B. (2018). *The Concept of Intelligence from Ancient Greek Philosophy to the Theory of Multiple Intelligences*. https://www.academia.edu/36815479/The_Concept_of_Intelligence_from_Ancient_Greek_Philosophy_to_the_Theory_of_Multiple_Intelligences

Grote, T., & Berens, P. (2020). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, *46*(3), 205–211. https://doi.org/10.1136/medethics-2019-105586

Haenlein, & Kaplan. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *ResearchGate*. https://doi.org/10.1177/0008125619864925

Halper, F. (2025). *TDWI Best Practices Report | Creating an AI-Ready Organization*. TDWI. https://tdwi.org/research/2025/03/adv-all-best-practices-report-creating-an-ai-ready-organization.aspx

Han, X., Zhang, Z., Ding, N., Gu, Y., Liu, X., Huo, Y., Qiu, J., Yao, Y., Zhang, A., Zhang, L., Han, W., Huang, M., Jin, Q., Lan, Y., Liu, Y., Liu, Z., Lu, Z., Qiu, X., Song, R., … Zhu, J. (2021). *Pre-Trained Models: Past, Present and Future* (No. arXiv:2106.07139). arXiv. https://doi.org/10.48550/arXiv.2106.07139

HHMGlobal, C. T. (2024, August 26). *The integration of AI in healthcare administration and management*. https://www.hhmglobal.com/knowledge-bank/news/the-integration-of-ai-in-healthcare-administration-and-management

Hintze, A. (2016, November 14). *Understanding the four types of AI, from reactive robots to self-aware beings*. The Conversation. http://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616

IBM. (2023, August 24). *What Is Sentiment Analysis? | IBM*. https://www.ibm.com/think/topics/sentiment-analysis

IBM. (2024, December 20). *AI Skills Gap | IBM*. https://www.ibm.com/think/insights/ai-skills-gap

Jandal, H. (2024, April 1). *AI-Generated Content: Balancing Automation & Authenticity*.

https://www.husamjandal.com/ai-generated-content/

Keskin, B. (2020). AI in operations management: Applications, challenges and opportunities. *Journal of Data, Information and Management*. https://www.academia.edu/48701447/AI_in_operations_management_applications_challenges_and_opportunities

Kirilenko, A. A., Kyle, A. S., Samadi, M., & Tuzun, T. (2011). The Flash Crash: The Impact of High Frequency Trading on an Electronic Market. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1686004

Klaise, J., Looveren, A. V., Cox, C., Vacanti, G., & Coca, A. (2020). *Monitoring and explainability of models in production* (No. arXiv:2007.06299). arXiv. https://doi.org/10.48550/arXiv.2007.06299

Krishnan, C., & Mariappan, J. (2024). “The AI Revolution in E-Commerce: Personalization and Predictive Analytics.” In L. Gaur & A. Abraham (Eds.), *Role of Explainable Artificial Intelligence in E-Commerce* (pp. 53–64). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-55615-9_4

Lin, B. (2025, January 6). How Are Companies Using AI Agents? Here’s a Look at Five Early Users of the Bots. *Wall Street Journal*. https://www.wsj.com/articles/how-are-companies-using-ai-agents-heres-a-look-at-five-early-users-of-the-bots-26f87845

Lincoln, Y. S., & Guba, E. G. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Program Evaluation*, *1986*(30), 73–84. https://doi.org/10.1002/ev.1427

Liul, M. (2022). *Custom AI Software vs Off-the-shelf Artificial Intelligence Solution*. Integrio

Systems. https://integrio.net/blog/custom-vs-of-the-shelf-artificial-intelligence

Loh, Beauchene, Lukic, & Shenoy. (2024, September 30). *Five Must-Haves for Effective AI Upskilling*. BCG Global. https://www.bcg.com/publications/2024/five-must-haves-for-ai-upskilling

Lucinity. (2024, November 6). *Reducing False Positives in Transaction Monitoring with AI - Transform FinCrime Operations & Investigations with AI*. https://lucinity.com/blog/understanding-false-positives-in-transaction-monitoring-what-causes-them-and-how-can-ai-can-reduce-operational-waste

Lumenalta. (2024, November 22). *Ethical considerations of AI: Fairness, transparency, and frameworks | Future of responsible AI*. Lumenalta. https://lumenalta.com/insights/ethical-considerations-of-ai

Marr, B. (2024). *Why Hybrid AI Is The Next Big Thing In Tech*. Forbes. https://www.forbes.com/sites/bernardmarr/2024/10/02/why-hybrid-ai-is-the-next-big-thing-in-tech/

McKinnon, T. (2018, October 31). *How Sephora Used Chatbots to Compliment In-store Service*. Indigo9 Digital Inc. https://www.indigo9digital.com/blog/learn-how-sephora-is-using-chatbots-to-create-a-better-customer-experience

McKinsey. (2022). *The state of AI in 2022—And a half decade in review | McKinsey*. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review

McKinsey. (2024). *Digital and AI leaders outcompete | McKinsey*. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/rewired-and-runnin g-ahead-digital-and-ai-leaders-are-leaving-the-rest-behind?

Morello, P. (2024, August 12). *How Amazon Uses Real-Time Data And Dynamic Pricing to Maximize Profits—Pricefy Blog*. https://www.pricefy.io/articles/amazon-real-time-data-dynamic-pricing

Morgan, L. (2022). *Enterprise hybrid AI use is poised to grow | TechTarget*. Search Enterprise AI. https://www.techtarget.com/searchenterpriseai/feature/Enterprise-hybrid-AI-use-is-poised-to-grow

Musani, P. (2023). *Decking the aisles with data: How Walmart's AI-powered inventory system brightens the holidays*. Decking the Aisles with Data: How Walmart's AI-Powered Inventory System Brightens the Holidays. https://tech.walmart.com/content/walmart-global-tech/en_us/blog/post/walmarts-ai-powered-inventory-system-brightens-the-holidays.html

Nayak, B. S., & Walton, N. (2024). History and Rise of Artificial Intelligence. In B. S. Nayak & N. Walton (Eds.), *Political Economy of Artificial Intelligence: Critical Reflections on Big Data Market, Economic Development and Data Society* (pp. 1–17). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-62308-0_1

Nguyen, T. (2024). *The Easiest Way To Custom AI Development—Neurond AI*. https://www.neurond.com/blog/custom-ai-development

OECD. (2022, February 22). *OECD Framework for the Classification of AI systems*. OECD. https://www.oecd.org/en/publications/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en.html

OneAI. (2023). *Pre-Trained Models & Embedded AI: Business Game Changers*.

https://oneai.com/learn/embedded-ai-pre-trained-models-business

Piatetsky, G. (2014, May 9). Did Target Really Predict a Teen's Pregnancy? The Inside Story. *Machine Learning Times*. https://www.predictiveanalyticsworld.com/machinelearningtimes/target-really-predict-teens-pregnancy-inside-story/3566/

PixelFish Media. (2024, January 25). *Cost-Effectiveness of AI-Generated Art in Marketing: A Comparative Analysis*. https://pixelfishmedia.com/blog/ai-generated-art-in-marketing/

Rajuhegde. (2024, March 16). MLOps: Continuous Delivery and Automation Pipelines in Machine Learning. *Medium*. https://medium.com/@rajuhegde2006/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning-093cd6e09fb3

Rebello, L. (2024). *Predictive analytics: Preparing your supply chain to overcome disruptions*. https://www.maersk.com/insights/digitalisation/2024/07/09/predictive-analytics-preparing-your-supply-chain-to-overcome-disruptions

Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach, 4th US ed.* https://aima.cs.berkeley.edu/

Ryseff, Bruhl, & Newberry. (2024). *The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed: Avoiding the Anti-Patterns of AI*. RAND Corporation. https://doi.org/10.7249/RRA2680-1

Semsarpour, M. (2023, March 29). *Gartner AI Maturity Model*. Medium. https://medium.com/@mohsen.semsarpour/gartner-ai-maturity-model-2c01fab629b6

Sequeira, A., & Tsang, F. (2024). *Emerging Legal Terrain: IP Risks from AI's Role in Drug Discovery*. Fenwick.

https://www.fenwick.com/insights/publications/emerging-legal-terrain-ip-risks-from-ais-role-in-drug-discovery

Sheludko, M. (2023, December 13). *Ethical Aspects of Artificial Intelligence: AI Ethics Explanation*. Software Development Blog. https://lasoft.org/blog/ethical-aspects-of-artificial-intelligence-challenges-and-imperatives/

Siemens. (2023). *The easy way to visionary robotics* [fw_Converting]. Siemens.Com Global Website. https://www.siemens.com/global/en/products/automation/topic-areas/tia/future-topics/simatic-robotics-ai.html

Smith, R. (2022). *Sorry, DABUS. AI cannot be an inventor on a U.S. Patent | Perspectives | Reed Smith LLP*. https://www.reedsmith.com/en/perspectives/2022/08/sorry-dabus-ai-cannot-be-an-inventor-on-a-us-patent

Suleyman, M. (2018). *A major milestone for the treatment of eye disease*. Google DeepMind. https://deepmind.google/discover/blog/a-major-milestone-for-the-treatment-of-eye-disease/

Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2019). Digital Twin in Industry: State-of-the-Art. *IEEE Transactions on Industrial Informatics*, *15*(4), 2405–2415. https://doi.org/10.1109/TII.2018.2873186

Turner, A. (2024, March 1). AI Compliance and Regulation: What Financial Institutions Need to Know. *ABA Banking Journal*. https://bankingjournal.aba.com/2024/03/ai-compliance-and-regulation-what-financial-inst itutions-need-to-know/

VentureBeat. (2019, July 19). Why do 87% of data science projects never make it into production? *VentureBeat*. https://venturebeat.com/ai/why-do-87-of-data-science-projects-never-make-it-into-production/

Verma, H. (2024, November 8). Renaissance Technologies Fund: Pioneering the Quantitative Trading Revolution. *Hedge Think - Digital Meeting Place for Fund Managers and Investors*. https://www.hedgethink.com/renaissance-technologies-fund-pioneering-the-quantitative-trading-revolution/

Walker, J. (2023). *How do airline pricing algorithms work?* https://eng.tourismandsocietytt.com/publications-articles/how-do-airline-pricing-algorithms-work

Weber, M., Engert, Schaffer, & Weking. (2022). Organizational Capabilities for AI Implementation—Coping with Inscrutability and Data Dependency in AI. *ResearchGate*. https://doi.org/10.1007/s10796-022-10297-y

Wheeler, T. (2023). *The three challenges of AI regulation*. Brookings. https://www.brookings.edu/articles/the-three-challenges-of-ai-regulation/

Wrenn, K., & Sohn, E. (2021). *Change Management for Artificial Intelligence Adoption*. https://www.boozallen.com/insights/ai-research/change-management-for-artificial-intelligence-adoption.html

Yosifova. (2024, February 29). *Data Scientist Job Market 2024: Analysis, Trends, Opportunities*. 365 Data Science. https://365datascience.com/career-advice/data-scientist-job-market/

Zhang, Y., Ren, S., Liu, Y., & Si, S. (2017). A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. *Journal of Cleaner Production*, *142*, 626–641. https://doi.org/10.1016/j.jclepro.2016.07.123

APPENDIXES

Appendix 1. Semi-Structured Interview Guide

1. Introduction and Context
   1.1. Could you tell me a bit about your current role and how you became involved in AI-related projects?
   1.2. What kinds of AI systems or initiatives has your organization explored or implemented so far?
   1.3. How would you describe your organization's current stage of AI adoption or maturity? What about on a scale from 1-5? What makes you feel it's at that level?
   1.4. Which business areas or processes are most affected by AI in your organization? Why were these areas prioritized?
2. Organizational readiness
   2.1. What initially motivated your organization to start implementing AI projects? Were there any business problems trying to be solved or rather opportunities to address?
   2.2. How would you describe the level of leadership sponsorship for AI projects in your organization? Has this level of support changed over time?
   2.3. How did different groups - management, technical teams, end-users, customers (if applicable) react to new AI projects?
   2.4. What kinds of preparations or change-management actions were taken before or during implementation?
   2.5. How would you describe your organization's culture towards innovation and change? Did it help or hinder AI implementation?
   2.6. How did internal talent, skills and cross-functional collaboration influence how well AI projects were implemented?
3. Technical Enablers
   3.1. From your perspective, what technical factors are most important when working on new AI projects? Could you share an example where the technical setup helped or hindered the process?

3.2. Was project scaling into production accounted for when creating new projects?

3.3. How easily were new AI systems integrated with existing IT infrastructure?

3.4. Did your organization rely mainly on custom-built solutions, off-the-shelf tools, or hybrid approaches? What influenced that choice?

3.5. How do you ensure AI systems remain reliable over time - for example, by incorporating new data, retraining, or evaluating performance? If applicable, who is responsible for that process?

4. AI Maturity

4.1. How do you recognize when an AI pilot project is ready to move into production?

4.2. Can you describe a project that successfully scaled from pilot to production? What made the transition possible?

4.3. Have there been projects that failed to scale? If yes, what held them back?

4.4. Between organizational and technical factors, which tended to have a bigger impact on the success of AI projects?

4.5. How long does it typically take for an AI pilot to be scaled to production? Is this significantly different from other IT projects?

4.6. Looking back, were there gaps between the initial roadmap and how things turned out when implementing AI projects?

5. AI Project Success

5.1. How does your organization measure the value of AI deployment? Was that value easy to identify or quantify?

5.2. How aligned were the technical teams and business leaders on desired outcomes?

5.3. How does the perceived value or impact of AI projects evolve over time? For example, does value increase, stabilize, or diminish as the system matures?

6. Reflection and recommendations

6.1. Based on your experiences, what do you think organizations most often misunderstand about implementing or scaling AI?

6.2. If you could redo one of your projects what would you approach differently?

6.3. What advice would you give to another organization starting its journey toward large-scale AI adoption?

6.4. Is there anything else you'd like to add that we haven't discussed but you think is important for understanding AI success or failure?

Appendix 2. Codebook derived from interviews

Name	Files	References
AI Adoption & Use Cases	1	1
Accounts Payable Automation	1	1
Ai used to help scale	1	1
All in one solutions with AI	1	1
Backup Options when Tools fail	1	2
Better idea generation than brainstorming needed	1	1
Continuous improvement bad use case	1	1
Duration of projects	5	7
Duration of projects differ based on function	1	1
Some projects are being implemented quickly with experience	1	1
Testing takes the longest	2	2
Experimentation importance	3	4
Project prioritization	2	3
Fraud Prevention	1	1
Function specific - vertical	2	4
General Productivity	2	3
Gemini and similar tools	1	1
Simple tasks work great	1	1
Internal Agents	5	12

Name	Files	References
Make people focus on more important parts	4	4
Pilot projects	3	3
Making Projects work in limited scope	2	2
POC importance	3	4
Sometimes POCs are difficult to transition to pilots	1	1
Scaling is not hard if planning is done right	1	1
Security Applications	1	2
Software giants don't work together	1	1
Summarizer - listener	1	1
Unstructured data good for gen AI	1	1
Implementation Challenges	0	0
3rd party support	3	3
Additional resources for maintenance	3	3
Choose the right direction	3	7
Cloud Based solutions are better	1	1
Data readiness	5	7
Data Drift and maintenance	1	1
Hard to plan ahead technologies needed to connect etc	1	1
LLMs are not good with structured data	1	1
Models improving and are able to work with worse data	1	1
Duration dependant on data availability and accessibility	1	1

Name	Files	References
Fast changes require upkeep	2	2
Fast changing technology requires fast processes	1	1
Hard to prove big initiatives for Leadership	1	1
Hard to spot mistakes	2	3
IT infrastructure limitations	1	2
Labor Regulation and Compliance	4	8
Lack of capability	3	4
AI is not smart	1	1
Hallucinations	5	6
Manual Testing Required	1	1
Misunderstanding AI Capabilities	4	5
Not always built around scalability	1	1
Off the shelve solutions are easy to implement	2	2
Pilot project limitations	1	1
Pilot tests	1	1
Goals	1	1
Process complexity	3	5
Regulations	2	2
EU is harder than other countries	1	1
Security concerns	3	3
Scaling too difficult - expensive	1	1
Technical misalignment mid-projects	1	1
Technology part doesn't take long	1	1

Name	Files	References
Tool inefficiencies	1	1
Workforce anxiety	2	3
LLMs not as bad as other technologies	1	1
Organizational Capabilities	0	0
Agile methods work better	2	2
Build vs Buy strategy	1	2
Capability development	2	2
Champions lead the new idea development	3	6
Change Management	4	6
Less resistance than other changes	2	3
No sugar coating for change management	1	1
Stakeholder change management	1	2
Cross team collaboration	2	2
Human guidance	3	3
Younger people tend to adopt new technologies faster	1	1
Lack of talent	2	6
Not knowing how to work with AI	1	1
Leadership change help with push for AI	1	1
Leadership Enablement & Compliance	5	6
Maturity dependant	4	6
Organization alignment on goals and time	1	1
Organizations lead by old people that don't understand technology	1	1

Name	Files	References
People don't know capabilities	1	2
Process standardization	2	2
Push from the management	3	4
Top-down strategic influence	2	2
Trainings & Workshops	6	11
Value Realization & Impact	0	0
Better results with better data	1	1
Employee market driving demand for AI	1	1
Limitations of off-the-shelf tools	4	6
Model & System Drift	1	2
Model decay	1	1
Model evaluation	4	10
AB Testing to check model efficiency	2	3
Accuracy is most important	2	3
Can't exactly pinpoint model value	1	1
Continuous Evaluation	1	1
Trade offs for different results	2	2
Using LLM for evaluation	1	2
no ROI on projects give uncertainty for leadership	1	1
No short term value	4	4
Experience builds skills and knowledge	2	3
Old ML & RPA importance	3	3

Name	Files	References
Depending on a case, LLM work better	1	1
Operational goals & Financial Efficiency	1	3
Organizational part makes project not deliver	1	1
Productivity improvements	1	1
Shifting focus from ROI to learning	1	1
Value realization limitations	1	1