this: POLITECNICO DI TORINO

Master Degree course in Engineering and Management

Master Degree Thesis

AI Adoption in Small and Medium Enterprises: Barriers, New Tools, and Operational Guidelines

Supervisors
Prof. Grazia Sveva Ascione
Prof. Elettra D'Amico

Candidate
Alessandro Buccini

Academic Year 2025-2026

Acknowledgements

My deepest gratitude goes to Claudia, the vital support of my life and the inspiration behind everything I undertake. Through every challenge she has been my strength and my steadiness, the constant presence without whom neither this thesis nor the journey behind it would have been possible. To her I owe more than these pages could ever express.

Abstract

SMEs have today gained access to artificial intelligence through an as-a-service market consisting of language-model APIs, no-code tools, and verticalised applications, but they remain uncertain on what problems to tackle first, what tools to use, and how to do it in a sustainable way. This thesis seeks to leverage that accessibility to enable practical adoption. It begins by examining the historical barriers to entry for SMEs and describes how these have been dismantled, with inferences now roughly a thousand times cheaper than three years ago, but with contemporary entry barriers having moved away from technology to focus on the areas of skills and data. It then defines a categorisation of problems that an SME can tackle using AI into five groups, each ordered by an increasing entry barrier, where the dominant paradigm for each category is to buy or to bundle, not to build. These are then combined into a gated, six-phase playbook that proceeds through problem selection, readiness assessment, technology selection, pilot, integration, and management, and is explicitly designed around the stage where SME AI adoption commonly breaks down: the gap between problem identification and problem execution. This thesis contributes a categorisation of problems that SMEs can solve with AI ordered by entry barrier as well as a reusable playbook that non-technical business owners can employ. The proposed guidance does not depend on evidence from tested deployment of the framework and therefore one area of future work consists of validating this framework by testing on live deployments, while another is the integration of increasingly sophisticated agentic systems.

Contents

Chapter 1

Introduction: SMEs and the New Wave of AI

1.1 Context

The role of SMEs in the economic landscape and the recent evolution of AI tools "as-a-service" (LLMs accessible via API, no/low-code platforms, open-source models, vertical solutions with integrated AI components) have created an unprecedented opportunity for SME-level adoption.

1.2 Problem (real gap)

SMEs are increasingly exposed to a growing number of tools, platforms, models, and "copilots". There remains limited operational clarity on which business problems to address as priority, which tool families to use, and how to structure a sustainable adoption path.

1.3 Thesis objectives

This thesis translates the new availability of AI-as-a-service to create a practical route an SME can take. It seeks to identify what kind of business problem an SME should tackle first, based on its constraints and stage of maturity; describe what solution patterns and toolkit families work with each problem type, and the build-versus-buy thinking that is appropriate; and bring it all together into a phased, stage-gated recipe that moves a business from first use case to managed, live deployment. It is intended to be actionable, rather than simply illustrative, to deliver advice that an SME owner-manager without a specific AI team can use.

1.4 Research questions

The objectives are answered through three research questions which are each presented in separate chapters.

1. Which categories of problem should an SME prioritise for AI adoption, and what makes one category easier to enter than another? (Chapters 4 and 5)

2. Which solution patterns and tool families address each category, and how should an SME decide whether to build, buy, or assemble? (Chapter 6)

3. How can an SME structure a sustainable, staged adoption path, from selecting the first use case to monitoring a deployed solution? (Chapter 7)

1.5 Scope

The scope of this work is the SME, the threshold for which the European Union defines as fewer than 250 employees, adopting AI through the as-a-service market rather than developing foundation models. It covers the problem types that recur in the SME evidence, namely document and process automation, customer service, marketing and sales, knowledge management, and decision support, together with the packaged and lightly assembled solutions available for them. It draws on peer-reviewed papers where available and on institutional and industry reports for market and adoption statistics, and it is restricted to recommendations that are feasible with the technology currently available. Validating the playbook against actual, live SME deployments is left for future work.

1.6 Structure of the Thesis

The remainder of the thesis is structured as follows. Chapter 2 explains the history and basic function of AI, and shows that accessibility, not ability, is now the limiting factor. Chapter 3 characterises SMEs and the current state of their AI adoption. Chapter 4 covers the barriers to adoption and how they have lowered between 2020 and 2024. Chapter 5 resolves the problems an SME can address with AI into five data-supported categories. Chapter 6 classifies the solution patterns and tool families that correspond to each problem, and the logic that informs the build-versus-buy decision. Chapter 7 assembles these into a six-phase implementation playbook. The concluding chapter answers the three research questions and sets out the contribution, the limitations, and the directions for future work.

Chapter 2

A Brief History of AI: Major Breakthroughs and Core Tasks

This Chapter traces seven decades of artificial intelligence (AI) research to argue that the historical gap between AI's technical maturity and its commercial accessibility has finally closed, opening a window for small and medium enterprises (SMEs) that did not exist in previous waves of AI. Three threads run through the narrative: the recurrence of disillusionment cycles, the so-called "AI winters", and their institutional memory among managers active during those periods; the gradual shift from symbolic, rule-based systems to statistical, data-driven learning approaches that culminated in the 2012 deep learning breakthrough and the 2017 Transformer architecture; and the emergence of a "democratization stack" of APIs, open-source foundation models, and cost collapses that places enterprise-grade AI within reach of SMEs for the first time. The chapter unfolds in four movements. It opens with the origins of the field and the dominance of symbolic, rule-based AI, followed by the first commercial wave of expert systems and the structural limits that exposed (Section 2.1). It then recounts the two "AI winters" and the statistical revival that, by 2012, replaced hand-coded rules with learning from data, before turning to the major architectural breakthroughs, the deep-learning result on ImageNet, the Transformer, and the late-2022 generative-AI inflection. From there it organises the core AI tasks into a capability taxonomy (perceptual, predictive, generative, and agentic) and closes by examining the "democratization stack" of APIs, open-source models, and collapsing costs that finally brings these capabilities within reach of even the smallest firms.

2.1 A Brief History of AI

2.1.1 Origins: From the Turing Test to the Birth of Symbolic AI (1950s-1970s)

The origins of AI as a field of research date back to the essay "Computing Machinery and Intelligence," written by Alan Turing in 1950 and beginning with the famous question "Can machines think?" [1]. Instead of attempting to answer this question through a philosophical definition, Turing proposed reformulating the problem in empirical terms. As he wrote in the first part of the article, the new form of the problem could be described "in terms of a game, which we call the 'imitation game'" [1]. In this game, a human interrogator must determine, through written questions alone, whether the responding entity is a person or a machine. Replacing a metaphysical question about thought with an empirical behavioral test provided a practical criterion that would define the discipline for decades and remains a touchstone in current discussions of machine thought.

The institutional origin of AI as a field of study occurred five years later. A two-month research project, scheduled to take place the following summer at Dartmouth College, was funded thanks to a proposal submitted by the field's four principal founders-John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon-on August 31, 1955 [2]. The proposal coined the very term "artificial intelligence" and articulated the field's founding aspiration in ways that have since become conventional: the investigation would proceed "on the conjecture that every aspect of learning or any other property of intelligence can in principle be described so precisely that a machine can be constructed to simulate it" [2]. Significantly, the original 1955 proposal already included "Neural Networks" as one of the seven topics [2], so even at the field's inception, this idea of neural networks was on the agenda alongside symbolic AI - the top-down approach that encodes human knowledge as explicit symbols and hand-written rules and reasons by manipulating them, rather than learning from data. The 1956 Dartmouth Conference brought together not only the four organizers but also several pioneers of computer programming, including Allen Newell, Arthur Samuel, Oliver Selfridge, and Herbert Simon [3]. After Dartmouth, the organizational structure of AI research began to be distributed primarily at MIT under Minsky, at Stanford University around McCarthy, and at Carnegie Mellon University, under the guidance of Newell and Simon [3].

The literature describes this dominant symbolic or representational approach of the early era in these words: "It was top-down and involved the use of natural language concepts (facts and rules) to represent and solve decision-making problems" [4]. The same source explains that "in a few decades (10, 20, 30, 40, now already 50 years) it would have been possible to synthesize cognitive processes and achieve general intelligence in machines" [4], and it is precisely this bold hypothesis that guided early AI research. The contrasting connectionist tradition, which would emphasize bottom-up learning from data, did exist as an alternative-the 1955 Dartmouth proposal itself listed "Neuron Nets" among its research topics, as noted above [2]-but it remained marginal throughout the first three decades of the field, while the symbolic paradigm shaped both research priorities and the construction of the first commercially relevant AI systems examined in section 2.1.2.

In the context of SMEs adoption of AI, it is important to emphasize the idea that the study of artificial intelligence has been strongly influenced, from its origins, by decision science management. According to the historical record, the term "artificial intelligence" was perhaps first used publicly at an IRE Symposium on the impact of computers on science and society in 1956 during a discussion, where the relationship between operations research and new artificial intelligence techniques can be observed in its embryonic stage

[3]. The Logic Theorist, a program that Newell, Shaw, and Simon completed in December 1955-which was the first example of a heuristic programming system-was developed at the Carnegie Mellon Graduate School of Industrial Administration, where interest in applications to business decision-making and cognitive psychology came and went in waves. As the same source notes, citing Simon himself, a primary reason why operations research had diverged from AI in its early stages was its inability to accommodate problem-solving strategies that more closely resembled human processes [3]. Therefore, from the very beginning, AI was not a spin-off of an academic discipline oblivious to the needs of organizations; it arose, in part, from the realization that classical optimization methods could not solve the complex, unstructured problems that managers and decision-makers faced.

These years of grandiose proclamations and early experiments with programs that could prove theorems, play games, and make limited use of natural language constitute the preparatory period for the next step: expert systems and the first commercial boom in AI, discussed in the next section.

2.1.2 Expert Systems: The First Commercial Wave and Its Structural Limits (1970s-1980s)

In the 1970s, AI evolved from a basic research program into a product with identifiable commercial applications. According to the survey literature, the field of expert systems began at Stanford in the mid-1960s with DENDRAL, a program designed to infer the structure of complex molecules from mass spectrograms at a level rivalling human experts [5]. Here lies the decisive break with the first symbolic wave. Where that earlier work pursued *general* reasoning through heuristic search procedures, expert systems made a *specific* expert's domain knowledge explicit and used it - rather than the search strategy - as the engine of the program. As the survey observes, DENDRAL's success showed for the first time that an intelligent program owed its power to what it knew about a narrow problem, not to the algorithmic search that had dominated previous efforts [5]. This insight - encode the knowledge of a human expert, not a general problem-solving method - propelled the entire expert-systems movement.

By the late 1970s and into the 1980s, expert systems had expanded from research prototypes to actual commercial systems. The commercial potential of the technology was exemplified by three classic systems: MYCIN, PROSPECTOR, and XCON [5]. The most frequently cited business application is XCON, the expert system implemented by Digital Equipment Corporation to assist in the configuration of complex computer hardware orders. As the survey points out, the annual savings for Digital Equipment Corporation resulting from the use of XCON were estimated at $20 million, and PROSPECTOR's contribution to the discovery of a $100 million molybdenum deposit is well documented. These were not experimental studies; they were economic achievements, and they stimulated widespread industrial interest in the technology. The survey states that over two-thirds of the Fortune 1000 companies applied expert systems to their daily business operations during the 1980s. At the time of that 1996 review, approximately 2,500 expert systems had been identified, suggested to represent only about twenty percent of the total, indicating that approximately 12,500 expert systems had been built worldwide.

Their primary focus was diagnosis, with the sectors seeing the greatest adoption being business, manufacturing, and medicine [5].

Nevertheless, though their success grew over time, it became evident that there were constraints on what the paradigm could achieve. In hindsight of 50 years, the literature admits to advancement in conceptual modelling, formalisation, and technique, but denies the stronger thesis that intelligence itself can be achieved by the symbolic manipulation of rules [4]. Early ambitions in AI were, on this view, unrealistic in that they failed to acknowledge the fundamental divide between human knowledge and the restricted, more specific knowledge that could be stored in an electronic computer [4]. For the purposes of the managerial argument, the result was simple: the thing that allowed expert systems to function, explicit, manually created rules, also constituted an ultimate limitation that simply more data would not surpass. The system's performance was determined by the rules the knowledge engineers had implemented, and if one wanted to apply it to another sector, the experts would once again have to be consulted and new rules entered.

This gap, quantified in various ways in the survey literature [5] regarding the commercial success of expert systems and conceptualized in the retrospective literature [4] regarding the structural limitations inherited from the symbolic systems paradigm, paved the way for the second of what have been termed "AI winters," discussed in the next section.

2.1.3 The AI Winters and the Institutional Memory of Failure (1973-1993)

The history of AI is punctuated by long periods of stagnation, retraction, and funding cuts, collectively known as the AI winters. The first episode, conventionally dated to the mid-1970s, was triggered by a report on the state of AI research that Sir James Lighthill prepared in 1972 and published in 1973 at the request of the British Science Research Council [6]. Commissioned as one layperson's assessment after two months of reading the literature and interviewing practitioners, the report was meant to guide the Council's funding decisions - and its verdict was harsh. Looking back on a quarter-century of work, Lighthill found that researchers themselves confessed to deep disappointment with what had been achieved, and concluded that nothing produced so far had matched the impact originally promised. Crucially, he located the cause not in lack of effort but in a technical wall: the field had failed to appreciate the *combinatorial explosion*. As a system's repertoire of knowledge grew, the number of possible states it had to search through expanded so fast that techniques which worked on small "toy" problems became intractable at any realistic scale [6]. The gap between grandiose promises and this fundamental scaling limit, more than any single failure, is what turned official confidence into withdrawal.

The report's institutional consequences were negative and long-lasting. It triggered a massive loss of confidence in AI among the British academic establishment and its funding body, an erosion of trust that persisted for nearly a decade [6]. The mechanism was concrete: a negative official assessment led the Science Research Council to withdraw support, which forced researchers out of the field and attached a lasting stigma to the very label "artificial intelligence." Donald Michie, one of the British researchers most directly affected, recalled seventeen years later that Lighthill's intervention had forced a divorce between the cognitive-science and knowledge-engineering wings of the young field, and it was Michie who applied the phrase "AI winter" to the period in which those who stayed in Britain saw their research budgets drastically cut [7]. The stigma outlived the funding freeze: when official efforts to revive the field arrived in the 1980s, such as the Alvey Initiative, they remained so marked by Lighthill's legacy that they deliberately avoided the term "artificial intelligence," substituting the neologism Intelligent Knowledge-Based Systems (IKBS) [7].

It is clear from the main texts that the second AI winter was a more complicated story than a mere slump. Michie noted that the funding cuts forced many British researchers out of the field [7]; but, according to a 1996 *IEEE Expert* survey, although the literature on expert systems had still been growing by roughly sixteen percent a year since 1988, the claim that the technology's death had arrived was judged overstated [5]. The two findings can be harmonised in this way: funding for general-purpose AI research went away, while the number of installations of specialised expert systems kept on increasing, as one subset of industrial software. What failed was not the technology so much as the institutions - the belief in the symbolic paradigm as a route to the complete intelligence its founders had promised, and, with it, the readiness to use the name "artificial intelligence". It is worth stressing this for what follows: even through the winter, statistical and data-driven methodologies were quietly maturing, thus paving the way for the revival of the 1990s and 2000s.

The following section traces the gradual emergence from this second winter, with the resurgence of statistical approaches and a reorientation toward empirical techniques that would lead, in 2012, to what is now referred to as the discontinuity of the Deep Learning Revolution.

2.1.4 The Quiet Revival: Statistical Machine Learning and the Road to 2012

The escape from the second winter for artificial intelligence was neither sudden nor spectacular. This process began in the 1990s and continued through the 2000s thanks to a shift in approach that was as quiet as it was profound: symbolic and rule-based systems were gradually replaced by statistical and data-driven methods that learned their behavior from examples rather than explicit programming. A crucial empirical milestone occurred in 2012, when a deep convolutional network trained by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton at the University of Toronto competed in the ImageNet Large Scale Visual Recognition Competition and achieved a result that was unexpected by the computer vision community [8]. As the researchers describe in their own paper, the network was trained to classify 1.2 million high-resolution images into 1,000 different classes, and the architecture they proposed consisted of 60 million parameters and 650,000 neurons organized into five convolutional layers and three fully connected layers. When evaluated in the 2012 ImageNet competition, the model achieved a top-five error rate of 15.3%, against the 26.2% of the second-best entry (Figure 2.1) [8]. This was not a marginal improvement but an absolute reduction in error of over ten percentage points.

The lesson the community drew was conceptual, not merely numerical: the winning system did not rely on visual features hand-engineered by experts, the standard practice until then, but *learned* those features directly from the raw pixels. A general-purpose learning procedure, given enough data and computation, had beaten decades of carefully crafted, domain-specific feature engineering - and it had done so on a hard, standardised benchmark. That single result convinced the field that deep learning was no longer a theoretical curiosity but a working technology.

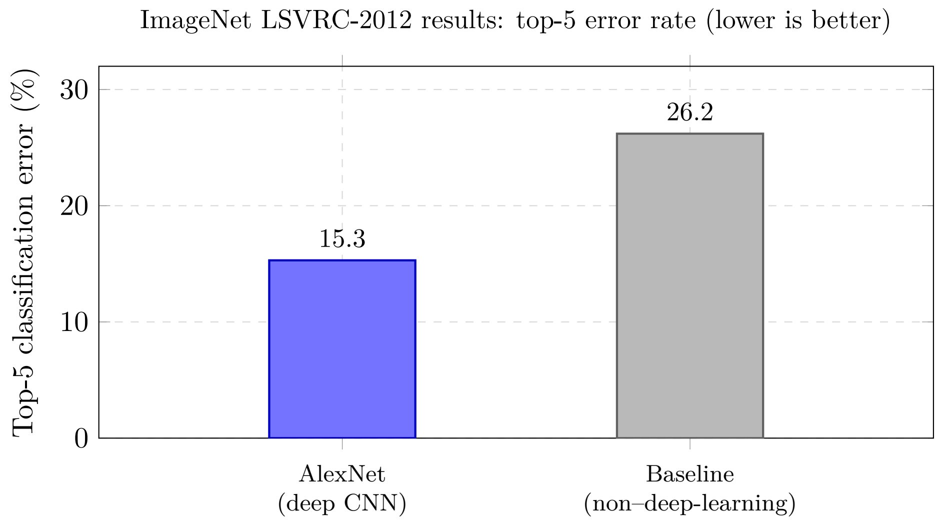


Figure 2.1. Top-5 classification error on the ImageNet LSVRC-2012 benchmark. AlexNet's 15.3% error was nearly 11 percentage points lower than the best competing approach (26.2%), the largest single-year jump ever recorded on the benchmark. A single result convinced the computer-vision community that deep learning was no longer a theoretical curiosity. Source: Krizhevsky, Sutskever & Hinton 2012 [8].

In the final line of their paper, the authors note that "[o]ur results can be improved simply by waiting for faster GPUs and larger datasets to become available" [8], a prophecy that would prove true for the following decade and foreshadow the cost and scalability curves that would eventually make modern AI accessible to organizations far smaller than the laboratories that developed it.

The breakthrough demonstrated in 2012 by Krizhevsky et al [8]-ten points better than the next best model in terms of error rate and running on a standard "off-the-shelf" GPU after just a few days of training-marks the starting point of the era of modern deep learning. So it was then that the "more data than knowledge" approach of statistical machine learning, as identified in the retrospective literature [4], finally gave way to true deep learning-of which AlexNet is the canonical empirical milestone-the moment of its first measurable, replicable progress on a standardized benchmark: the kind of thing that, during the winter (or winters) of AI, managers had learned to dismiss as premature. And this turning point in 2012 ushers in the following macro-section of this chapter, dedicated to the major breakthroughs in the era of deep learning and its architectural progress, with the transformer, which has since made possible, today, generative artificial intelligence systems capable of reshaping the horizon of possibilities for SMEs.

2.2 Major Breakthroughs

This Section examines the three modern milestones that, together, define the technical foundation of contemporary AI and explain why the technology has only recently become operationally accessible for SMEs: the 2012 AlexNet result on ImageNet [8], the 2017 Transformer architecture [9], and the late-2022 inflection driven by consumer-accessible generative AI [10]. Figure 2.2 situates these three breakthroughs within the longer arc of the discipline, alongside the foundational moments of §2.1 (Turing 1950, Dartmouth 1956) and the two "AI winters" that shaped the institutional memory of today's business decision-makers. The remainder of this section discusses each milestone in turn, from the paradigm shift away from hand-coded rules (§2.2.1) to the Transformer era (§2.2.2) and the generative-AI inflection point (§2.2.3).

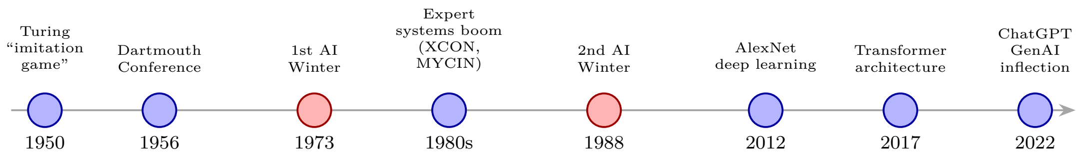


Figure 2.2. Seven decades of AI history at a glance: scientific milestones (blue) alternate with the two "AI winters" (red) that shaped the institutional memory of today's business decision-makers. Sources: Turing 1950 [1]; Dartmouth proposal [2, 3]; Lighthill report [6]; Michie retrospective [7]; expert-systems survey [5]; AlexNet [8]; Transformer [9]; Vinsel on the 2022 inflection [10].

2.2.1 The Paradigm Shift: From Hand-Coded Rules to Learning from Data

The key intellectual shift that distinguishes the current wave of AI from the first wave of AI is learning behavior from data, rather than encoding it into an explicit set of rules. Having concluded the previous section with AlexNet [8] as the main empirical result of this shift, the following section revisits the conceptual reasons before returning to the architectural ones. The latest consensus review on machine learning as the transition was underway is a 2015 paper published in Science, which defines machine learning as the discipline that "addresses the question of how to build computers that improve automatically through experience" and places it "at the intersection of computer science and statistics, and at the center of artificial intelligence and data science" [11].

According to that review, this transition had taken place in the two decades that came before the analysis, as machine learning had achieved tremendous success over that period, turning from a laboratory curiosity into a practical tool for commercial use [11]. The major factors responsible were the drive to develop new technologies, algorithms, and theories, together with the continuous explosion in the volume of low-cost data accessible on the Internet. The review finds that by 2015 the widespread adoption of such techniques had become pervasive across all of science, technology, and commerce, leading to more evidence-based decision-making in a wide range of activities, from healthcare and manufacturing to education, financial modeling, policing, and marketing [11]. This is significant for the managerial argument of this thesis: by the mid-2010s, machine learning was already a functional technology used in business decision-making, not an area of experimental research.

In a single practical observation, that review captures the essence of the shift: for many applications, it is much easier to train a system with examples of the desired input-output behaviour than to hand-code the correct response to every possible input [11]. Programming by specification was being superseded by programming by example.

At this point, LeCun, Bengio, and Hinton, in their 2015 review, emphasize that traditional machine learning methods were limited in their ability to process raw natural data [12]. For decades, creating a pattern recognition or machine learning system required the design of a feature extractor, where the correct selection of features was a problem of domain expertise and careful engineering in order to transform the data into a more meaningful internal representation, or a feature vector from which a machine learning component-often a classifier-could detect or classify patterns in the input. The deep learning approach has replaced this process of manual feature design with a procedure in which features are learned from the data: as LeCun and his co-authors put it, these feature hierarchies are not designed by human engineers but learned from the data using a general-purpose learning procedure [12]. The elimination of the need for human-based feature engineering has enabled the industrial-scale application of machine learning to messy and unstructured data-such as images, audio, and natural language text-which is the type of data for which previous rule-based symbolic systems were ill-suited.

2.2.2 The Deep Learning Era and the Transformer (2012-2017)

Following AlexNet in 2012 (as discussed in Section 2.1.4), the subsequent five years represented the most productive five years in terms of architecture and experimentation in the entire history of artificial intelligence. In 2015, LeCun, Bengio, and Hinton published their paper that provides an explicit list of the consumer-visible applications already powered by machine-learning systems by 2015: "from web searches to content filtering on social networks to recommendations on e-commerce websites," as well as the identification of objects in images, the transcription of speech to text, and the matching of news items, posts, and products to user interests [12]. In other words, in just three years since the breakthrough of AlexNet, deep learning had transformed from a research gimmick into an industrial-scale platform for consumer Internet products-something that previous waves in the field had repeatedly failed to deliver.

LeCun and his co-authors conclude their review with a prescient comment. "We believe that deep learning will achieve much more in the future, because it requires little manual engineering. It can therefore more easily benefit from increases in computing power and data," they say. The pace will only accelerate with the arrival of new techniques and algorithms being developed, even as we speak, in the next generation of deep neural networks [12]. It was an architecture that demonstrated more than any other that this prediction was correct. This architecture, known as Transformer [9], was proposed in 2017 by a team of researchers primarily from Google Brain and Google Research, two years later.

Vaswani et al. present "a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely" [9]. The benefits were twofold. On the standard machine-translation benchmarks of the time, the Transformer set new state-of-the-art results: 28.4 BLEU for the WMT 2014 English-to-German task, "improving upon the best published results, including ensembles, by over 2 BLEU points," and 41.0 BLEU for the WMT 2014 English-to-French task [9]. Second, and more importantly in hindsight, the new state-of-the-art was achieved at a fraction of the computational cost of the best previous models; for the English-to-French result, "only 3.5 days of training on 8 GPUs" were used, as the authors explicitly state, "which represents a small fraction of the training costs of the best models in the literature." [9]

And in fact, by replacing recurrent computation-the de facto standard for sequence modeling up to that point-with the parallelizable attention mechanism, the Transformer "enables significantly greater parallelization and can achieve a new state-of-the-art in translation quality after just 12 hours of training with eight P100 GPUs" [9]. To put it another way, the same architecture not only produced better translation performance but also reduced training time by an order of magnitude, and was explicitly designed to scale efficiently on parallel hardware. It is also noteworthy that the Transformer originated from an industrial laboratory: seven of the eight co-authors of the 2017 paper were working at Google Brain or Google Research at the time of writing [9]. The point is not that corporate labs employ better researchers than universities - talent exists in both places - but that, increasingly, the computational resources and large-scale data needed to produce and validate such architectures sit inside a handful of very large technology firms. This concentration matters directly for SMEs: it means the most capable AI tools tend to reach the market as commercial products and cloud services from those same firms, rather than as freely reproducible academic results - a dependence revisited in Section 2.4.1.

2.2.3 Generative AI and the 2022 Inflection Point

The previous subsection followed the architectural lineage from AlexNet, which demonstrated the power of deep convolutional networks in 2012 [8], to the Transformer, which introduced the attention-based architecture in 2017 [9]. Within that lineage, the launch of ChatGPT in November 2022 marks a turning point - but it is essential to be precise about what kind. It was *not* a technical revolution: the underlying architecture was already five to ten years old, and no fundamentally new model was unveiled. What changed was access. By contrast, ChatGPT was an *adoption* inflection point, the moment a capability that had previously existed only in research labs and engineering teams became usable, through a simple conversational interface, by non-technical business users. This subsection therefore treats it as such, focusing on how the significant shift was in the relationship between AI and business decision-makers, not in the science.

According to a trend analysis [10], what distinguishes generative AI, from a technical standpoint, is that it consists of "machine learning models, such as ChatGPT, Bing AI,

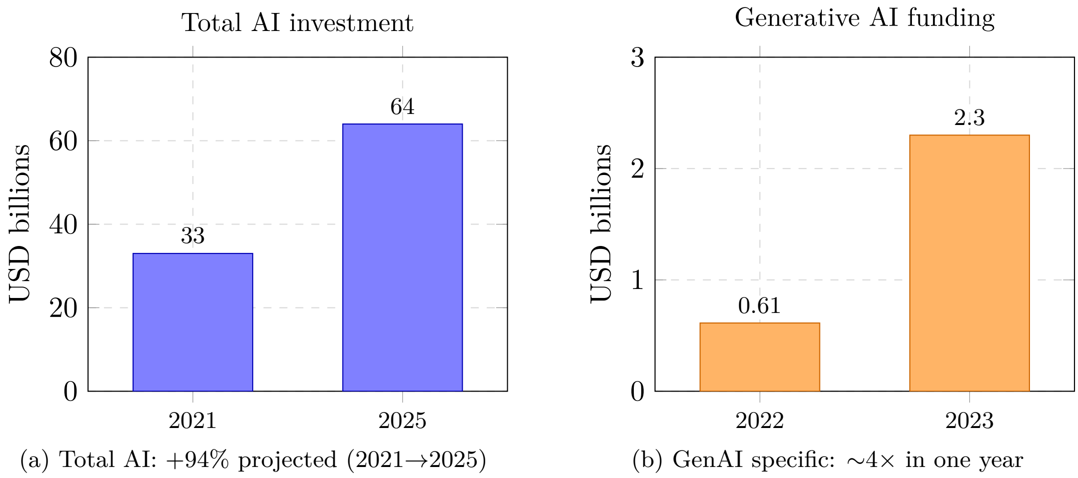


Figure 2.3. The financial signal of the 2022 inflection. (a) Total AI investment was projected to nearly double between 2021 ($33B) and 2025 ($64B). (b) Generative-AI-specific funding alone jumped roughly fourfold in a single year, from $613M in 2022 to $2.3B in 2023, the financial hallmark of a tipping-point event. Source: Vinsel 2023, MIT Sloan Management Review [10].

DALL-E, and Midjourney, trained on massive databases of text and images to generate new text and images based on a prompt." The scale of the capital that flowed toward these systems is itself a marker of the inflection: as Figure 2.3 shows, generative-AI-specific funding jumped roughly fourfold in a single year while total AI investment was projected to nearly double over 2021–2025. The list is particularly significant because it also compares ChatGPT with similar products launched by Microsoft and OpenAI (and Stability AI / Midjourney; it is important to note that the context is November 2022, when these tools first became available to the general, non-technical public), and thus the list contextualizes the moment when these various text and image generation systems became publicly available. The innovation was not the neural network architecture itself, as shown in the previous subsection; this had been in circulation in research at least since the AlexNet breakthrough in 2012 [8] and the 2017 Transformer paper [9]; rather, it was the convergent arrival of user interfaces for consumer-accessible models, which had reached a quality where the results were good enough to allow business users to actually utilize the output.

2.3 Core AI Tasks and Capabilities

Having traced how AI became technically capable, this section now classifies that capability, for the purpose of SME adoption, into four key families: perceptual, predictive, generative, and agentic AI. The key details of these families are summarised in Table 2.4, drawn from a survey of the machine-learning and deep-learning literature and of recent trend analyses [10–13]. The four subsections that follow give further explanation of each family in turn, with an emphasis on the concrete business tasks it can perform.

Perceptual AI Vision, speech, language understanding *Quality inspection, OCR, transcription, email triage*	**Predictive AI** Classification, forecasting, recommendation *Credit risk, demand forecast, churn, predictive maintenance*
Generative AI Text, image and code synthesis *Marketing copy, mockups, code drafting, customer-service replies*	**Agentic AI** Multi-step task automation *Tool use, function calling, cross-system orchestration*

Figure 2.4. A capability taxonomy of the four families of AI most relevant to SME adoption, classified by *what each can do* - perceiving inputs, predicting outcomes, generating content, and acting autonomously - rather than by when it appeared. The families are complementary and frequently combined within a single application; each corresponds to one of the four subsections below. Sources: LeCun, Bengio & Hinton 2015 [12]; Jordan & Mitchell 2015 [11]; Vinsel 2023 [10]; Davenport & Bean 2026 [13].

2.3.1 Perceptual AI: What Machines Can See, Hear, and Read

Among the most mature and widely commercialized capabilities are those grouped under the category of "perception": systems that perceive the world by acquiring visual, auditory, and textual inputs, performing on par with humans-and in many specific cases, better than humans.

LeCun, Bengio, and Hinton showed just how far machine perception enabled by deep learning had advanced by the mid-2010s, powering web search, social-media filtering, and e-commerce recommendations, as well as increasingly common consumer devices such as cameras and smartphones [12]. In other words, they note that "machine learning is used to identify objects within images, convert speech to text, match news articles, posts, or products to readers' preferences, and provide relevant search results" [12]. In a very real sense, image classification, text-to-speech and speech-to-text conversion, text content matching, and search result ranking are all examples of perceptual capabilities: the system is fed unstructured data-whether a photograph, an audio stream, or a search term-and returns structured data, in the form of a label or transcription, or an ordered list of documents.

In a business context, quality inspection on a production line, transcription of customer service calls, classification of incoming emails based on their underlying intent, or the structured extraction of information from scanned invoices are all examples of perceptual artificial intelligence applied to business processes that were once performed manually by people. The literature states that by 2015, "the use of this data-intensive method had spread 'across science, technology, and commerce,' with applications including 'healthcare, manufacturing, education, financial modeling, policing, and marketing'" [11].

2.3.2 Predictive AI: Classification, Forecasting, and Personalization

A second family of capabilities goes beyond perception to prediction: systems that apply learned historical patterns to categorize new observations, predict future states, or curate personalized recommendations for individual users. Among the many types of AI in production, classification-which assigns data points to one of several categories-is the most common and is used to predict credit risk, identify transactional fraud, diagnose patients from medical images, and predict customer churn [11]. Regression, where the target is a continuous value, is used for demand forecasting, price optimization, and predictive maintenance in manufacturing [11]. Recommendation systems, which use both regression and classification to determine user interests, are a major source of revenue in e-commerce and online content [12].

Predictive analytics is likely the first AI technology that small and medium-sized enterprises will implement, as the technology can utilize existing structured data-such as sales databases, customer lists, and inventory trackers-rather than requiring a substantial investment in unstructured data storage infrastructure, as might be necessary for perceptual and generative AI.

2.3.3 Generative AI: Content, Code, and Creative Synthesis

The third category of capabilities-and the best known among those that emerged since late 2022-is generative AI: tools capable of generating new content, rather than categorizing information or making predictions based on pre-existing data. As the trend analysis describes it, generative AI is a type of "machine learning model (think ChatGPT, Bing AI, DALL-E, and Midjourney, among others) that has been trained on massive databases of text and images to generate new text and images in response to a prompt" [10].

The real-world applications of generative AI can be organized into three categories of modalities [10]. Text generation is used in marketing copy, email drafting, report summaries, translation, and customer service bot responses. With image generation, it is possible to transform descriptions into images, create mockups, produce images for advertising, or develop drafts for projects. Code generation, which likely attracts the least public interest but the most attention from companies, can help programmers and software engineers write, debug, and refactor code at a much faster pace, while also allowing those without a technical background to create small-scale automations via text commands. The economic weight of this modality is not incidental: software engineering is among the four business functions identified by an industry analysis as capturing the majority of generative AI's estimated value [14].

From a managerial perspective, these tools may increase output per employee: controlled studies report a 15% productivity gain for support agents [15] and a 40% reduction in task time, with higher quality, on writing tasks [16], while an industry analysis estimates generative AI could automate activities absorbing 60–70% of employees' time [14] - though transforming this into output growth without added headcount depends on redeploying the freed time.

2.3.4 Automation and Decision Support: AI in Business Processes

In addition to perception, prediction, and generation, there is a fourth group of capabilities that directly impact business operations: they automate sequences of actions that were previously part of the human experience. This cluster also refers to the shift from classic robotic process automation (RPA)-that is, logic-based scripts replicating human inputs within structured software applications-to an "analytical, generative, agentive" approach to AI integration, as a recent trends review [13] calls it.

The latest trend in this family is the rise of agentic AI, a new class of systems capable of completing complex, multi-step tasks without constant human guidance, using tools, invoking functions, and making decisions along the way. As noted in a 2026 review of major AI trends, published in the MIT Sloan Management Review, agentic AI is "the most hyped trend since, well, generative AI," while generative AI is now "in Gartner's trough of disillusionment" [13]. This is a balanced critique; "the errors agents make now are too numerous to trust them with managing business processes of substantial monetary value," but the review holds that "in about five years, AI agents should be able to handle most transactions in many large-scale business processes" [13].

2.3.5 Latest developments

The 2026 trends review is structured to look beyond what lies ahead and focuses on five observations that directly address the issue of adoption by SMEs [13]. First, the review predicts that the bubble will burst, citing Amara's Law, by which "we often overestimate the short-term impact of new technologies while underestimating their long-term impact" [13].

Second, the review describes the emergence of what it calls "AI factories"-enterprise systems that bring together existing technology platforms, methodologies, data, and algorithms to enable the development of AI systems in a "fast and easy" manner [13]. Third, it notes, by way of example, that generative AI is becoming less and less a tool for individual productivity and more and more an organizational resource. Alternatively, it suggests that generative AI should be used in strategic, enterprise-level use cases, such as supply chain management, research and development, and sales functions-not to "speed up the creation of a blog post" [13].

Fourth, and as noted earlier, the review considers agentic AI to be overhyped in the short term but predictably useful over the next five years, provided that current limitations in error rates, cybersecurity, and alignment with human values can be reduced [13].

Fifth, there is no established AI governance structure: the 2026 survey reports that 39% of large companies have a Chief AI Officer (CAIO), "but there is little consensus on whom this role reports to" [13].

2.4 From Breakthroughs to Business Applicability

2.4.1 The Democratization Stack: Open-Source Models, APIs, and the Cost Collapse

The capabilities listed in the previous section-perceptual, predictive, generative, and agentic-do not, in and of themselves, explain why small and medium-sized enterprises (SMEs) should consider adopting artificial intelligence now rather than in five or 10 years. The difference between, say, 2012-when deep learning really took off-and now is not the capabilities themselves, but rather how those capabilities are made accessible. The purpose of this section is to describe how recent structural changes in technology have made AI economically and technically accessible to organizations that do not have the same research budgets, engineering staff, and computing power as the companies where this technology was initially developed.

Several developments, which roughly parallel each other between 2020 and 2025, have collectively lowered the barrier to entry. The first are large language models, available as paid APIs, such as those developed by OpenAI (the GPT family), Anthropic (Claude), Google (Gemini), and a growing number of other providers. These services transform what was previously a capital investment in specialized hardware and engineering expertise into a running operational cost priced per token, eliminating the need to build and train the model or maintain GPUs in-house. The second key factor is the launch of high-quality open-source foundation models, such as Meta's eponymous LLaMA product family, the Mistral models from a French startup of the same name, and similar offerings from other organizations, which allow for on-premises hosting, customization on proprietary datasets, and use completely independent of costly licensing fees. Finally, rounding out the picture is the growing competition among providers. In this regard, as noted in the MIT Sloan Management Review, "[. . .] competition [among providers] has caused a sharp increase in AI spending, which is projected to rise from '$33 billion in 2021 to $64 billion in 2025' and 'generative AI funding [...] will rise from $613 million in 2022 to $2.3 billion in 2023." [10] The practical consequence of this competition is a sustained and profound reduction in the cost per unit of AI capacity.

A 2015 scientific review predicted that today's advances in machine learning, in addition to the development of better machine learning algorithms themselves, were made possible by "the ongoing explosion in the availability of online data and low-cost computing" [11]. What could not have been predicted is how quickly this explosion would translate into commercial prices: by 2025, an SME can obtain state-of-the-art access to a language model via an API for a fraction of a cent per query - something that would have seemed unlikely at the time of writing. Taken together, these three factors - easy-to-use APIs, a vibrant open-source ecosystem, and collapsing costs - constitute what is here termed the democratization stack. This is the crux of the chapter, the fulcrum on which the rest of the thesis turns: for the first time in the seventy-year history of AI, the binding constraint on adoption is no longer the *capability* of the technology, which large firms had long enjoyed, but its *accessibility*, and that constraint has now fallen away even for the smallest businesses. The historical gap between technical maturity and commercial reach, the thread running through this entire chapter, has finally closed.

2.4.2 The Vendor and Platform Ecosystem

The democratization stack is not an isolated phenomenon. It is mediated by a rapidly expanding ecosystem of vendors and platforms that encapsulate AI capabilities into products accessible to the non-technical user. For the SME operator, the most important question is not "how does a transformer work?" or "What tools can I buy, deploy, or subscribe to this quarter?"

The AI ecosystem can be mapped on a continuum ranging from general-purpose assistants to highly specialised embedded systems. The trend analysis lists general-purpose products such as "ChatGPT, Bing AI, DALL · E, and Midjourney" [10], to which can be added Microsoft 365 Copilot, Google Workspace with Gemini integration, and Anthropic Claude for Work. These solutions will integrate AI tools that provide assistance with text drafting, data analysis, and code generation, as well as content summarization, operating within the software suites that most SMEs use on a daily basis. At the middle layer, no-code and low-code automation platforms, such as n8n, Make (formerly Integromat), and Zapier's AI agent feature, enable business users to create complex multi-step workflows to connect AI models with databases, CRMs, and communication tools-all without writing code. At the other end of the spectrum, vertical-specific SaaS platforms are integrating AI into the very core of the product, such as AI-powered forecasting in CRMs, AI-driven categorization in accounting software, and AI-powered chatbots serving as customer support agents.

A further recent development is the class of agentic AI platforms, which perform multi-step tasks on the user's behalf. As the trends review notes, some companies have begun to "develop specific trusted agents that can be reused across various departments within the organization and test cross-organizational agents with cooperative suppliers or customers" [13]. These systems represent a paradigm shift in the conception of AI: rather than assisting human activity, they act as agents that carry out that activity themselves. More broadly, this layering of products and services on top of foundation models is part of a wider movement from discrete AI *models* to AI *platforms*, a process of platformisation in which a handful of providers supply the underlying capability and a surrounding ecosystem of vendors packages it for end users [17]. A complementary literature frames the same trend as the rise of "AI-as-a-service", in which AI is consumed on demand through cloud interfaces rather than built in-house [18]. Other scholars are more circumspect, cautioning that such platform dependence reinforces the power asymmetries and lock-in risks already familiar from the broader platform economy, an exposure especially acute for resource-constrained SMEs [19].

2.4.3 Risk Landscape for Businesses

The fact that today's AI tools are widely available does not mean that the potential dangers of their use have disappeared. These risks are particularly relevant for small and medium-sized enterprises, which typically lack dedicated compliance, legal, and cybersecurity functions, and tend to be overlooked in the early, frenetic stages of adoption.

Based on a systematic review of 63 peer-reviewed studies, the literature argues that "numerous pilot projects for generative AI have suboptimal results, primarily due to a lack of alignment with an organization's strategy and a lack of adequate governance structures": a problem that underscores the need to view AI risks as organizational issues rather than merely technical tasks [20]. The most discussed topic in the context of generative AI is hallucination: fluent and confident text output that may nevertheless be factually incorrect. This is a problem that every SME will face if it chooses to use AI for anything from drafting customer communications and legal opinions to writing technical specifications. Another type of risk is bias: since they are trained on Internet-scale data, models can acquire and amplify biases present in the data, potentially leading to discriminatory hiring recommendations, credit assessments, or biased customer interactions. The third concern is intellectual property exposure-specifically, the possibility that proprietary company data fed into a cloud-based AI model could be retained, used for further training, or shared with other users of the same service.

A specific risk for European SMEs, however, is linked to regulations such as the EU AI Act (Regulation 2024/1689), effective as of August 2024, which introduced a classification system for artificial intelligence systems based on their level of risk for those deploying them in EU member states [21]. Although most SME use cases, such as text generation and internal analysis, fall into the lowest risk categories, systems that impact hiring decisions, creditworthiness, or safety-critical processes may be considered high-risk and therefore subject to compliance assessment, documentation, and monitoring requirements. However, no matter how proportionate the burden may be, it entails an organizational cost that must be considered by SMEs.

A fifth threat-and one that is frequently underestimated-consists of what experts refer to as invisible artificial intelligence: an ad hoc adoption of AI, managed in a disorganized manner by individual employees outside the company's approved technology stack. In a review of studies, it is observed that "weak organizational support" and "limited awareness of AI applications" are obstacles to the implementation of a systematic approach [20]. According to the same review, "financial constraints, limited digital skills, and weak organizational support" are regularly cited in the literature as challenges to the implementation of a systematic approach [20]. In other words, without formal policies to guide them, employees end up using a plethora of tools of their own choosing, which results in untracked data and unregulated decision-making processes. It is not the use of AI by employees that should be feared-often, at the individual level, its use increases productivity-but rather the fact of not knowing how AI is being used, what data is being shared with third-party applications, and how decisions are influenced by AI-generated outputs.

2.4.4 Conclusion

This chapter has traced the history of artificial intelligence from Turing's 1950 formulation of the imitation game [1], through the commercial optimism and subsequent disillusionment of the expert systems era [5], the two AI winters and their enduring institutional legacy [6, 7], the quiet statistical revival that culminated in the deep learning breakthrough of 2012 [8], the Transformer architecture that enabled large language models [9], and the generative AI inflection of late 2022 that brought these capabilities to the attention of the business mainstream [10]. Thus, it has identified the main groups of AI capabilities, including perceptual, predictive, generative, and agentic ones [11–13], and described the structural changes related to costs, access, and the ecosystem that have made these capabilities economically and technologically accessible for the first time even to small and medium-sized enterprises [10, 11].

This chapter has also identified two underlying themes that will be revisited and discussed throughout the remainder of this thesis. The first is the institutionalized memory of failure from AI's previous winters: the skepticism documented in sections 2.1.3 and 2.2.3, which drives managers to use AI tools with a caution that is historically rooted rather than technically grounded [4, 6, 7]. Next is the disparity between technological readiness and operational readiness: the findings of the systematic review literature, which indicate that "the success of [technology] adoption depends on organizational readiness, managerial expertise, technological infrastructure, and environmental support" [20], most of which SMEs currently lack.

Chapter 3

SMEs and AI Adoption

3.1 Characteristics of SMEs

In the European Union's widely used definition, an SME is an enterprise with fewer than 250 employees; within this category, a small enterprise has under 50 employees (and under €10 million annual turnover) and a micro enterprise under 10 employees (and under €2 million turnover)[22]. Despite variations at the margins, the core idea is that SMEs operate on a smaller scale than large corporations, which has important implications for their structure and behavior. Other jurisdictions define SMEs differently [23], but the EU definition provides a useful benchmark for this thesis. Small and medium-sized enterprises (SMEs) are generally defined by their modest scale in terms of employees and financial turnover.

In Europe, for example, micro, small and medium enterprises together employ about two-thirds of the business sector workforce and generate over half of value-added[24]. However, SMEs tend to lag behind larger firms in productivity, with labor productivity on average only about half that of large companies [24]. Their prevalence is even higher in emerging economies, where SMEs can constitute almost four-fifths of employment [24]. They make up over 90% of all firms globally, contribute about 50% of GDP, and account for the majority of employment in most countries [25]. Closing this gap is a major economic opportunity, as one analysis estimates that raising SME productivity to best-practice levels could be equivalent to a boost of 5% of GDP in advanced economies (and 10% in emerging ones) [24]. This ubiquity means that the productivity and innovation capacity of SMEs have large aggregate effects on economic growth and social welfare.

The characteristics of SMEs differ markedly from those of large enterprises [26, 27]. In most small businesses, the founder or owner-manager plays a pivotal role in decision-making and may wear multiple hats (strategy, sales, HR, etc.) due to the limited managerial layers. By nature of their size, SMEs typically have simpler organizational structures and fewer specialized resources. SMEs often operate with tight financial resources and limited access to capital, making large up-front investments (in technology, R&D, etc.) more challenging. This can allow for agility and quick decision cycles, but it also means that capacity is constrained, considering there are simply fewer people to dedicate to any given function. In other words, scale itself is a barrier – smaller firms do not benefit from the same economies of scale, whether in procurement, production, or data collection, that large firms enjoy [28]. They also generate and handle less data compared to big firms: a systematic review highlights data availability and data quality as recurring technological obstacles for SMEs adopting AI, noting that data quality is often more problematic for SMEs than for larger enterprises due to missing data collection standards and data cleaning practices [29].

Some SMEs are high-growth startups at the cutting edge of innovation, while many others are traditional, family-run businesses embedded in local markets [30]. It's important to recognize this diversity because generalizations about SMEs may not apply uniformly. SMEs are an extremely heterogeneous group. Their sizes span a wide range: even within the "SME" bracket, a medium-sized firm of 200 employees has vastly different management structures and market reach than a micro-enterprise of 2–3 staff. They include everything from a solo self-employed tradesperson or micro-retailer, to a 20-person local manufacturer, up to a 200-person software company. They operate across all sectors of the economy – manufacturing, services, retail, agriculture, and more – and their business models vary widely. For instance, micro-firms (those under 10 employees) make up the vast majority of all businesses (over 90% of firms in most OECD countries are micro[28]), yet standard business statistics often exclude them, meaning the overall picture of "SME" trends can be skewed toward the upper end of the size range. Overall, though, most SMEs share the constraint of limited scale and resources, which influences their approach to technology adoption, risk management, and innovation.

While a smaller size can confer flexibility and close customer relationships, it also often means weaker internal capabilities in areas such as dedicated IT staff, data infrastructure, and cash buffers for investment. They are typically more resource-constrained but agile relative to large firms. These traits have a direct bearing on how SMEs adopt new technologies like AI. As the following sections explore, these characteristics shape both the needs and the challenges SMEs face in implementing AI solutions. In summary, key characteristics of SMEs include their small scale, broad prevalence in the economy, and resource constraints (financial, human, and data).

For the purposes of this thesis, *artificial intelligence* (AI) refers to the family of techniques surveyed in Chapter 2 — primarily statistical machine learning approaches that learn patterns from data, generative models based on the Transformer architecture, and rule-based automation augmented with learned components. AI is considered here as it is encountered by SMEs in practice: as managed cloud APIs (e.g., large-language-model endpoints from major providers), as open-source models that can be self-hosted, or as functionality embedded into vertical SaaS tools that SMEs already use. This operational definition deliberately excludes purely symbolic, hand-coded "expert systems" of the kind discussed in Chapter 2, which are no longer the dominant paradigm for the tools relevant to SME adoption today.

3.1.1 Key Business Processes in SMEs

Whether a company is large or small, it needs to develop products or services, reach customers, manage finances, handle procurement and supply chains, and administer its operations. Despite their size, SMEs must perform essentially the same fundamental business functions as any larger enterprise. However, the way these processes are executed in SMEs can differ notably from large firms, often being less formalized or supported by fewer and simpler systems. This section outlines the key business processes in SMEs and how SMEs typically manage them, setting the stage for understanding where digitalisation and AI can intervene.

Nonetheless, SMEs engage in planning , such as deciding what products or services to offer, which markets to target, and how to allocate their limited resources. Indeed, only a minority of SMEs currently leverage advanced data analytics for strategic planning – for example, EU data indicate relatively low uptake of Big Data analysis among small firms [31].

Strategic Management and Planning: In an SME, strategic planning and high-level management are usually concentrated in the hands of the owner or a small leadership team. This function is intimately tied to the owner's expertise and experience. Decision-making tends to be intuitive or based on personal knowledge of the market, rather than driven by large-scale data analysis. There may not be separate strategy departments or extensive analytics guiding long-term plans. In practice, this means SME strategic decisions are often made with limited formal data support, which can be a handicap in complex or rapidly changing environments. A mapping of SME business functions to ICT use shows that activities like "Direction and strategic planning" have among the lowest adoption of digital tools (such as business intelligence or enterprise planning software) by small companies [31].

In SMEs, these back-office processes are typically handled by a very small team (or even a single person) and are often only partly digitized [28].

Administration and Finance: All businesses must handle administrative tasks – book-keeping, invoicing, regulatory compliance, HR administration, IT support, etc. Interestingly, small firms have not lagged far behind larger ones in certain basic admin digitalization: for instance, the prevalence of using the internet for dealings with public authorities (e.g. One notable aspect is interaction with government (for taxes, social security, permits, etc.), which has increasingly moved online. Many small firms rely on off-the-shelf accounting software and external accountants for annual financial statements. Day-to-day, spreadsheets and paper files might still be in use alongside basic accounting packages. e-filing taxes) is quite high across all firm sizes[31]. There is little difference between small, medium and large enterprises in the incidence of online administrative interactions (B2G) and electronic invoicing, as these have become widespread practices[31]. Indeed, only 38% of small enterprises in the EU use ERP systems, compared to 86% of large enterprises[32]. That said, SMEs still typically have less integrated internal IT systems for administration. The vast majority of small businesses therefore manage administration through a patchwork of simpler tools, which can lead to duplicated data entry and inefficiencies (a point revisited in Section 2.3). They may not use enterprise resource planning (ERP) software to unify finance, HR, and operations data in one system, because ERP implementation requires a "critical size" and significant investment that many small firms are not ready for[31].

SMEs in different sectors will have very different operational processes – a small manufacturer manages factories, inventory, and suppliers, whereas a small consultancy's "production" is delivering expertise. Generally, SMEs have fewer layers of specialization in operations management. For example, in manufacturing SMEs, machine tools and production lines might not be instrumented with the Internet of Things or advanced sensors due to cost, meaning data on production performance is limited. Many operational workflows are manual or only semi-automated in smaller firms. A single person might be in charge of purchasing, managing inventory, and overseeing production scheduling. Supply chain management (SCM) software adoption is also lower in SMEs than in large firms [31]. Operations, Production and Supply Chain: This core process involves how the SME creates its product or delivers its service, including procurement of inputs and logistics. Therefore, while SMEs carry out all the same operational steps (sourcing raw materials, producing goods, delivering to customers), they often do so with less automation and data visibility across the chain. The gap in adoption between small and large firms widens as technologies become more sophisticated – "large firms have invested more intensively in the integration of their business processes (ERP, CRM, SCM), tools for strategic planning, and tools for production and logistics management"[31]. Instead of sophisticated SCM systems that automatically share data with suppliers or optimize logistics, small firms often coordinate via phone calls, emails, and Excel sheets. For instance, the use of RFID tracking in logistics or real-time inventory management systems is much less common in SMEs[31].

Web presence and social media are now standard even for very small firms: in 2023, for example, 78% of EU enterprises had a website [32], and this figure included the majority of small businesses (about 76% of small firms have a website, compared to 95% of large firms)[32]. In fact, diffusion data show that the adoption gap between small and large companies is smallest in areas like social media use for marketing [32]. Marketing, Sales and Customer Service: Reaching customers and generating sales revenue is a critical process for SMEs. Over the past decade, however, digital marketing tools have become both crucial and relatively accessible for SMEs. Traditionally, many small businesses relied on local word-of-mouth, repeat customers, and analogue methods (physical storefronts, print advertising). Social media use by SMEs is also widespread as a marketing channel, often because it is low-cost and easy to adopt. Small firms have eagerly embraced platforms like Facebook, Instagram, or local e-marketplaces to advertise and sell. While the overall percentages are lower, the gap here is not as huge as one might expect – even small retailers increasingly sell via online channels or platforms. In the EU, roughly 25% of small enterprises use Customer Relationship Management (CRM) software to manage client interactions, which is lower than the 60%+ of large firms that do so, but still a significant uptake [32]. This indicates that many SMEs handle customer relationships in relatively informal ways (spreadsheets or personal contact lists), but a growing share are adopting CRM tools to organise customer data and sales leads. E-commerce is another part of the sales process: around 22% of small EU firms received orders online (vs ~45% of large firms) according to recent statistics [31]. Overall, marketing and sales is often one of the first areas SMEs digitalize, because tools like websites, online marketplaces, and social media provide a relatively easy way to expand market reach[31]. However, SMEs may not exploit more advanced marketing analytics to the same degree as large firms; for instance, using AI-driven targeting or big data analysis of customer behavior is still mostly the domain of larger companies (the share of firms performing big data analysis is much higher among large enterprises)[31].

To summarise, SMEs perform the full range of business processes – strategic planning, administration, operations, and customer-facing activities – but often in a simpler, more people-intensive manner [28]. This pattern reflects the lower barriers and immediate benefits of front-office digital tools for small businesses, versus the higher complexity and cost of back-office automation. [31]. For example, the median adoption rates for social media and electronic sales among small firms are relatively close to those of larger firms, whereas adoption rates for integrated internal systems (like ERP for production or SCM for logistics) show a much bigger gap [31]. As the discussion turns to the challenges SMEs face, it will become clear how these process characteristics – fragmented systems, manual workarounds, and selective digital uptake – create structural issues that can impede effective use of technologies like AI.

3.2 Structural Challenges: Fragmented Information, Unstructured Data, Manual Work

SMEs' internal processes, as described above, often involve fragmented information systems, unstructured data, and manual workflows. These structural characteristics pose significant challenges when it comes to adopting advanced digital technologies, particularly the AI techniques defined in Section 3.1, which thrive on integrated, high-quality data and automated processes. This section examines these interrelated challenges in SMEs' operational context and their implications.

For example, if an SME lacks an ERP system (and recall, fewer than 4 in 10 small firms use ERP[31]), then financial data might live in an accounting spreadsheet, sales leads in an email inbox or a CRM tool, and inventory records in yet another file. These systems often don't talk to each other. Because many SMEs do not use enterprise-wide software platforms, their information tends to be siloed across different tools or departments. But for a vast number of small businesses, data silos and duplication are daily reality. The result is that getting a "single source of truth" about the business – say, a consolidated view of sales, costs, stock, and customer feedback – is laborious. One major issue is fragmentation of information. Someone (often the owner-manager) has to manually pull information from various places to make decisions. As firms grow to the upper end of SME range, they start encountering the complexity that compels investing in integrated systems [32]. When data is fragmented, it becomes difficult to apply AI tools that typically require a comprehensive dataset or real-time data streams covering multiple aspects of operations. This fragmentation not only reduces efficiency (time is wasted switching between systems and reconciling data) but can also lead to inconsistencies and errors in information.

Closely related is the challenge of unstructured or low-quality data in SMEs [29]. Even when SMEs use digital tools, the data may remain inaccessible for advanced analysis – for instance, using separate Excel files for each month's sales, instead of a continuously updated database. Small businesses may generate far less data than a large company, and what they do generate is often not collected or stored in a structured format. Research confirms that SMEs often suffer from limited data availability and quality relative to large firms[29]. Or a small service provider might have customer information in the form of email correspondence and paper notes rather than a structured database. An AI system is only as good as the data fed into it; if an SME's data is scattered (fragmented) and largely unstructured (text documents, images, free-form emails), significant preparation and cleaning are required before any AI can be applied. In fact, "lack of data" is frequently cited as a barrier when SMEs consider implementing AI, because many AI algorithms require large, well-organized datasets to train on[33]. This puts them at a disadvantage – while a large corporate might instantly leverage its centralized data warehouse for an AI pilot, a small firm first has to find and clean its data, which can be a daunting task. SMEs typically do not have dedicated data engineering teams to undertake that work.

Due to limited automation, employees in small businesses spend a considerable amount of time on routine, repetitive tasks that could in theory be automated by digital systems. Another structural challenge is the prevalence of manual work and processes in SMEs. Administrative burden is one clear example. A recent survey in Germany quantified this burden: the average SME spends around 7% of its working hours on administrative processes, equating to 32 hours per month per business on paperwork and compliance [34]. Half of SMEs spent 15 hours or less, but a substantial minority spent far more[34]. SMEs often lack specialized compliance or administrative departments, so tasks like filling out government forms, processing invoices, or compiling reports fall on the owner or a small team. Furthermore, SMEs perceive bureaucracy as a top impediment – in that German survey, SMEs rated bureaucracy as the single greatest risk to their competitiveness in their domestic market[34]. While some administrative tasks will always exist, larger firms can assign staff or use enterprise software to handle them more efficiently, whereas a small firm's key people must divert their attention to these manual chores. The lack of digital solutions (such as integrated e-government interfaces or automated reporting tools) exacerbates this. This illustrates how manual effort on non-productive tasks (tax forms, documentation, regulation compliance) drains SMEs' limited time.

Manual work is not limited to administration. Therefore, many such processes remain manual. But in an SME, the up-front cost and complexity of automation can be prohibitive. In daily operations, SMEs often have people performing tasks that, in a more digitalized setting, could be done by software or machines. This has a cumulative effect: processes that are manual are typically slower and generate less data, which then ties back to the earlier point about limited data for AI. For instance, taking orders and transcribing them into an invoice, scheduling staff rotas by hand, or manually inspecting production outputs for quality – these are all tasks that can be automated or augmented by AI (through e-commerce systems, workforce management software, or machine vision for quality control). It's a kind of vicious cycle – because processes are manual and not instrumented, there is little data to analyze; because there's little data and no existing automation, it's harder to introduce AI to improve the process.

GPUs). Inadequate infrastructure (such as older computers, no high-speed network, lack of cloud computing resources) can hinder the deployment of modern AI solutions which often require substantial computing power or specialized hardware (e.g. Without a stable and modern IT backbone, even trialing AI tools (which might involve, say, installing

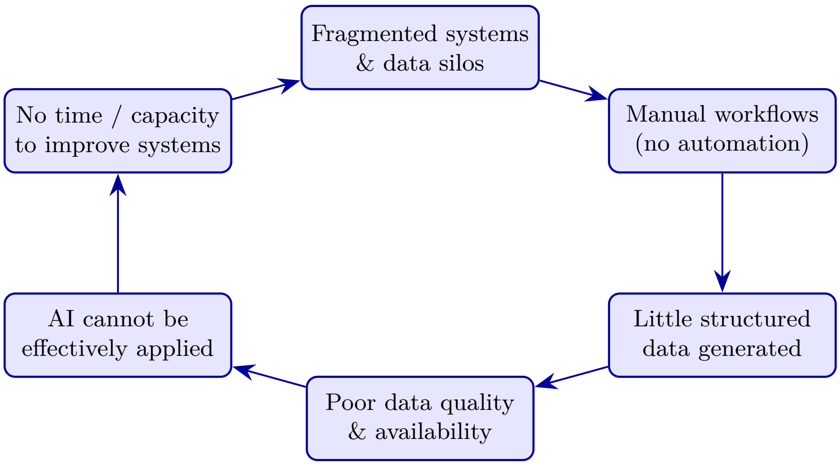


Figure 3.1. The structural "vicious cycle" that keeps many SMEs locked out of effective AI adoption: fragmented systems generate little structured data, manual workflows reduce data capture further, poor data quality makes AI ineffective, and the resulting strain on staff leaves no capacity to improve the systems. Breaking the cycle typically requires external support from technology providers, government programs or cloud-based AI-as-a-service offerings [35, 36].

new software or connecting to cloud AI APIs) becomes difficult. Another dimension of the structural challenge is the limited IT infrastructure and expertise within SMEs. In a systematic review of AI in SMEs, inadequate infrastructure was identified as one of the most common barriers to implementation, alongside lack of knowledge and high costs [33]. Many small companies do not have an in-house IT department or dedicated IT professionals – they might rely on external IT support for setting up computers and networks, and use basic cloud services for email and file storage. This means SMEs might lack access to scalable computing resources that are useful for AI workloads. Larger firms are reported to be "more proactive in externalising the development and maintenance of their IT systems" via cloud services, whereas smaller firms often stick with in-house, ad-hoc systems[31]. Many SMEs also have not fully migrated to cloud-based systems – although this is changing, smaller firms have historically been slower to adopt cloud computing for their core business processes [31].

Finally, it's worth noting that these structural issues feed into each other. Fragmented systems and data lead to manual reconciliation, manual processes produce less structured data and consume staff time, and limited time and IT expertise make it harder to improve the systems. This underscores why external support (from technology providers, government programs, etc.) is often deemed necessary to help SMEs break out of this pattern [35, 36]. It can become a trap where an SME is "too busy to improve" – the day-to-day operational strain leaves little bandwidth to implement transformative digital changes.

These factors result in inefficiencies and missed opportunities even before considering AI. More critically for the purposes of this thesis, they create a less-than-ideal starting point for AI adoption. The next section will examine the current state of AI adoption in

SMEs, keeping in mind how these challenges translate into slower uptake and the need for targeted strategies. Overcoming these structural issues is therefore a prerequisite for – or at least a parallel task to – successful AI implementation in the SME sector. AI thrives on abundant, well-structured data and automated, well-defined processes – conditions that many SMEs currently do not meet. In summary, SMEs commonly face structural challenges in the form of non-integrated information systems, poor data quantity/quality, and heavy reliance on manual work.

3.3 State of AI Adoption in SMEs

Throughout this section, the term *artificial intelligence* is used in the operational sense introduced in Chapter 2: predictive machine learning applied to structured business data (e.g., demand forecasting, churn prediction), extbfgenerative AI based on large language models accessed via APIs (e.g., document drafting, customer-facing chatbots, knowledge-base Q&A), and rule-based automation augmented with AI components (e.g., robotic process automation in administrative workflows). The discussion below covers adoption across all these task families, with the caveat that public statistics rarely distinguish them at this level of granularity.

This section assesses how widely SMEs are adopting AI as of the mid-2020s, what types of AI applications they use, and how their adoption compares to larger firms. Given the structural barriers outlined, it is perhaps not surprising that small and medium enterprises have, until recently, been slow adopters of artificial intelligence. However, the landscape is rapidly evolving. The analysis draws on the latest available data from official statistics and surveys to present a nuanced picture: one of historically low uptake but a current trend of acceleration and narrowing gaps.

3.3.1 Overall Adoption Rates and Trends

As one study noted, around 90% of SMEs had no AI applications in use circa 2020 [33]. SMEs, with their resource constraints, largely stayed on the sidelines. This latent interest has begun translating into real adoption as AI technologies become more accessible (think of user-friendly cloud AI services and off-the-shelf AI solutions) and as competitive pressures mount. Until the early 2020s, AI adoption among SMEs was minimal. Nonetheless, awareness of AI's potential was growing – a 2021 survey found that 77% of respondents saw great possibilities for using AI in SMEs [33], indicating a positive mindset even if actual implementation lagged. This was a time when AI was mostly the domain of large tech companies and some early-adopting large enterprises. In other words, only a small single-digit percentage of small firms were utilizing any form of AI at the start of the decade.

As of 2025, 19.9% of EU enterprises employ at least one AI-powered technology [37][28]. By the mid-2020s, data show a notable uptick in SME AI adoption, especially in certain regions. This means roughly one in five companies (across all sizes) is now an AI user – more than double the share just two years earlier. But the latest data for 2024 and 2025 show a sharp increase. In 2021, only about 7.7% of EU enterprises (with 10+ employees) had adopted any AI technology; by 2023 this was around 8.1%, essentially flat[24]. The surge coincides with the proliferation of more accessible AI tools (for example, the widespread publicity of generative AI in 2023). It's worth noting that these figures cover firms with 10 or more employees; micro-enterprises are not included in EU surveys, so overall SME adoption (if micros were counted) would appear lower. In the European Union, the official statistics reveal a jump in the share of enterprises using AI. Nonetheless, the trajectory is clear: AI adoption is accelerating among smaller firms.

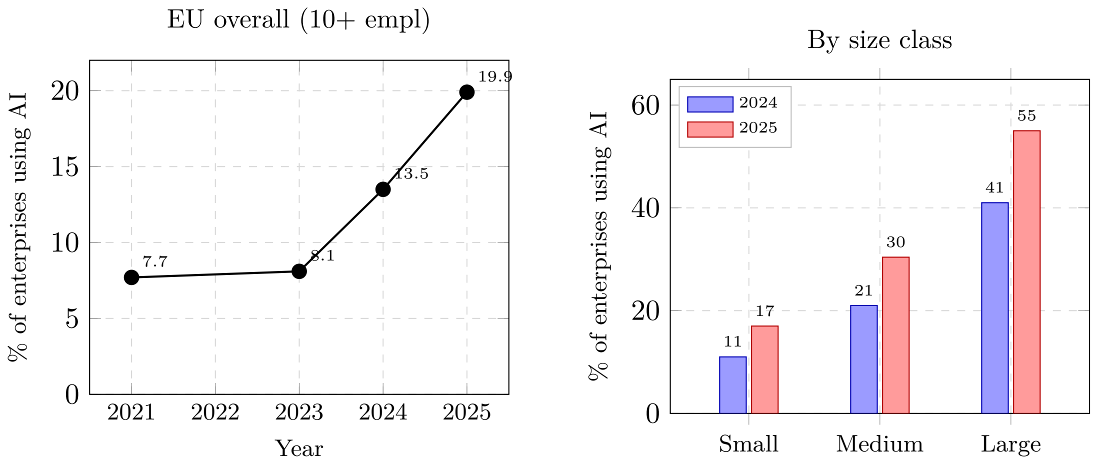
The data consistently show a size-based adoption gap. In other words, a majority of large firms are now leveraging AI in some form, but only about one in six small businesses are. Despite this progress, SMEs still notably trail large enterprises in AI uptake. Figure 3.2 illustrates this gap, and also how it evolved from 2024 to 2025 – all size classes saw increases, but large firms maintain a substantial lead. In the EU in 2025, while 19.9% of all enterprises use AI, the breakdown by size was 17.0% of small enterprises (10–49 employees) and 30.4% of medium enterprises (50–249 employees), versus 55.0% of large enterprises[28].

Small firms (<50 employees) saw an increase from 11% to 17% adopting AI between 2024 and 2025, but remain well behind large firms (>250 employees) where adoption grew from ~41% to 55%[28], as shown in Figure 3.2.

The gap is rooted in the factors discussed earlier (resources, skills, complexity), and indeed this difference can be "explained, for example, by the complexity of implementing AI technologies in an SME, economies of scale and cost issues"[28]. In fact, from 2024 to 2025 the share of small firms using AI jumped by 6 percentage points (from ~11% to 17%), whereas large firms' share increased by ~14 points (41% to 55%). Medium firms saw an increase of nearly 9 points. This indicates momentum is building among SMEs, although large companies, with their greater means, also accelerated their uptake. As shown above, the adoption gap in relative terms is significant – large enterprises are over three times as likely as small ones to be using AI. Encouragingly, the gap narrowed slightly in the latest year as small and medium firms' adoption rates grew faster (in absolute terms) than large firms'.

This has led to the suggestion that, at the current rate of acceleration in SME uptake, small businesses may close the gap with larger firms in overall AI adoption rates within approximately one year [38].

In the United States, at the same time, small businesses have historically lagged larger firms in adopting new technologies—a pattern often described as a "digital divide"—but the AI gap appears to be closing faster than in previous technology cycles. According to the Business Trends and Outlook Survey, as of late 2023 about 6.3% of small businesses (<250 employees) were using AI in their operations, compared to 11.1% of large businesses [37]. In other words, the rapid diffusion of certain AI tools is allowing smaller firms to catch up more quickly than they did with, say, internet usage in the 1990s. This is a reminder that "SME" is a broad category and early adopters can exist at any size. One fascinating data point: in the U.S. survey, the very smallest businesses (those with 1–4 employees) showed a higher rate of AI use than other small firms (5–249 employees), creating a U-shaped relationship where adoption is highest at the micro and large ends, and lowest in the middle[38]. This U-shape likely reflects the presence of tech-savvy startups among the micro firms – essentially, a subset of very small companies that are



(a) Flat 2021–2023, then surge post-GenAI release

(b) Adoption rises across all sizes, gap widens

Figure 3.2. Evolution of AI adoption among EU enterprises 2021–2025. (a) Overall trend for enterprises with 10+ employees: essentially flat between 2021 (7.7%) and 2023 (8.1%), then surge after the public release of generative AI tools, reaching 19.9% in 2025. (b) Breakdown by size class for the two years where separate figures are reported (2024 and 2025): the absolute gap between small and large firms widened (+14 p.p. for large vs +6 p.p. for small). All three size classes grew between 2024 and 2025, but the absolute gap between small and large firms widened (+14 p.p. for large, +9 p.p. for medium, +6 p.p. for small). Sources: Eurostat [24, 28].

born digital and adopt AI early, even as many slightly larger traditional small businesses are slower.

3.3.2 Variation by Sector and Country

Certain sectors lend themselves more to AI solutions and as a result have higher uptake [24]. By contrast, in more traditional sectors, AI use is still nascent. In construction, only about 10.8% of firms were using AI in 2025 [28]. AI adoption among SMEs is not uniform across industries or regions. According to the 2025 data [24], the information and communication sector (which includes software companies, IT services, etc.) had by far the highest share of firms using AI – about 62.5% of enterprises in that sector used some AI, a figure that includes many SME tech companies [28]. These are sectors where the product is often digital or knowledge-based, making AI tools immediately applicable (for example, AI for coding assistance in software firms, or AI for data analysis in consulting). Professional, scientific and technical services (which covers things like consulting firms, R&D services, engineering SMEs) also had a relatively high adoption rate at 40.4%[28]. Retail and manufacturing were in the middle of the pack (somewhere in the range of 15–25% of firms using AI)[28]. It suggests that when considering AI adoption, the sectoral context of SMEs matters: a small tech startup and a small construction contractor have very different incentives and opportunities to use AI. Similarly low adoption (well under 15%) was observed in sectors like transportation & storage, and accommodation/food services, which are dominated by SMEs that often operate on thin margins and simpler technology. The broad point is that AI is more relevant for certain activities and SMEs in those fields have taken it up more readily, whereas others have yet to find compelling use cases or face higher barriers.

Crucially, in all EU countries the trend from 2024 to 2025 was upward – even the lowest-adopting countries saw increases in the share of firms using

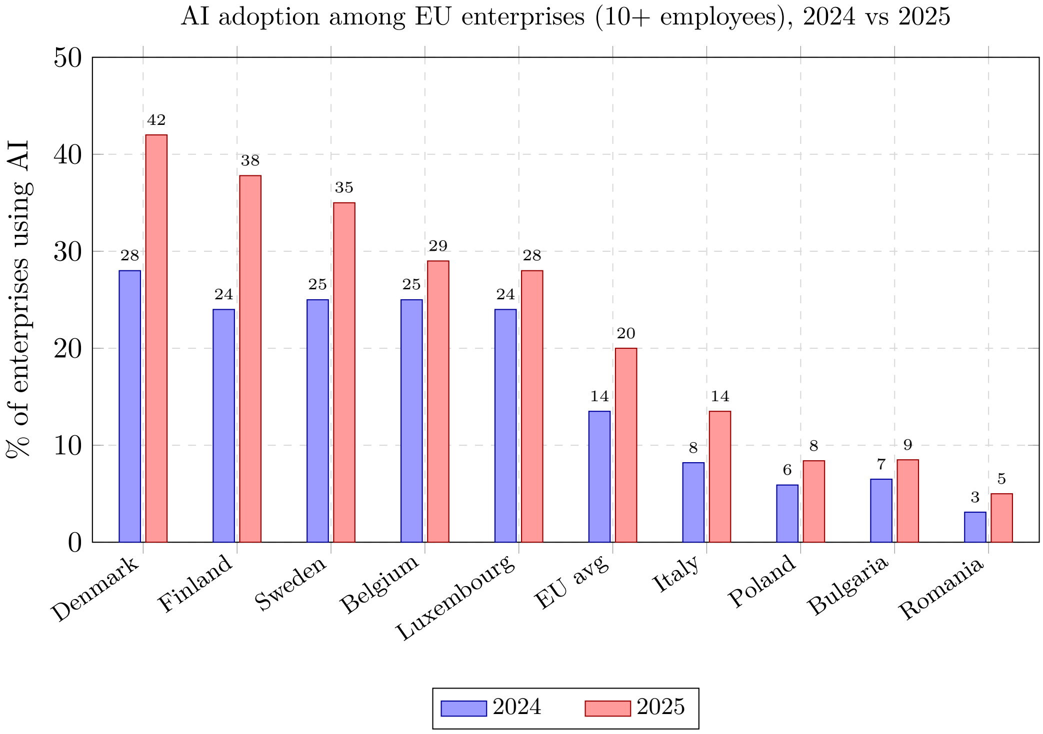


Figure 3.3. AI adoption growth among EU enterprises between 2024 and 2025, highlighting the divergence between Nordic frontrunners and Southern/Eastern European laggards. Italy sits below the EU average. Source: Eurostat [24].

AI[28]. There are also regional differences. Within Europe, adoption rates among SMEs vary by country. For instance, Denmark leads with about 42% of enterprises using AI, followed by Finland (~38%) and Sweden (~35%) [28]. On the other end, some Southern and Eastern European countries have much lower figures – Romania (~5%), Bulgaria (~8.5%), Poland (~8.4%) were the laggards in 2025[28]. Generally, countries with high digitalization and innovation indices (often in Northern and Western Europe) show higher SME AI usage. These differences reflect variations in factors like digital infrastructure, availability of AI talent, and industry mix, as well as policy support. Many of those enterprises are SMEs, given the structure of Nordic economies and their strong digital culture. This indicates a Europe-wide movement towards greater AI diffusion, albeit at different speeds.

3.3.3 How SMEs Are Using AI

The usage patterns of AI among small firms often center on a few key application areas, typically those that are easier to implement or promise quick wins. Looking beyond whether SMEs use AI or not, it's revealing to see what they use it for when they do. Surveys and data suggest that SMEs using AI tend to focus on customer-facing and analytical applications, rather than heavy automation of core production.

About 34.7% of AI-using enterprises employed AI for marketing or sales purposes (e.g. According to data on the purpose of AI systems in enterprises (note: this covers all sizes, though patterns can still be distinguished by size), the most common uses of AI in 2025 were in marketing and sales, and in business administration. AI-driven advertising, customer segmentation, recommendation engines), and 31.1% used AI for organising business administration or management processes (this could include AI in accounting software, scheduling, or decision support for management)[28]. This indicates that early AI adoption in firms – including SMEs – is concentrated in areas like improving customer acquisition or automating routine office tasks, rather than, say, running a warehouse with AI robots (which is far more complex and costly). On the other hand, very few firms used AI for logistics (only 6.1% of AI users applied it to logistics/supply chain)[28].

When broken down by enterprise size, there are notable differences in use cases. These gaps suggest that even when SMEs venture into AI, they often stick to less complex applications, potentially due to resource constraints or lack of necessary data in production processes. AI-based threat detection), whereas only 14.5% of small enterprises did so[28]. For example, among companies using AI, nearly 47.5% of large enterprises used AI for ICT security (e.g. Similarly, AI for streamlining administrative/management processes was used by 43% of large firms but only 28.5% of small firms that have AI[28]. Large firms are much more likely than small firms to use AI for certain purposes such as cybersecurity, production optimization, or management processes. In production and operations, 33.5% of large AI-using firms applied AI to production processes, compared to 19.0% of small AI-using firms[28]. Small firms appear to prioritize AI that can help with external growth (marketing/sales) or internal administration, rather than AI that requires overhauling how they make their core product.

Marketing automation was a standout example where small firms were ahead – many small companies are leveraging AI tools to automate marketing emails, social media posting, or customer follow-ups[38]. Interestingly, the same survey showed small businesses leading in several AI use cases. It found that small and large businesses that use AI actually employ a similar number of distinct AI use-cases on average – about 2.0 use cases for small firms versus 2.1 for large firms, implying that once a small firm is on board with AI, it can experiment across multiple areas almost as much as a big firm[38]. Of 17 AI application categories considered, small businesses led large businesses in adoption in almost half of them[38]. This is likely because such tools are readily available (often as affordable SaaS products) and do not require heavy IT integration. This pattern is consistent with U.S. survey evidence [38], which nonetheless identifies specific use cases where small businesses lag notably behind larger firms. Small firms were also quite on par in use of AI for things like text analysis or natural language processing in their operations. These are use cases that often demand more investment in integration and data

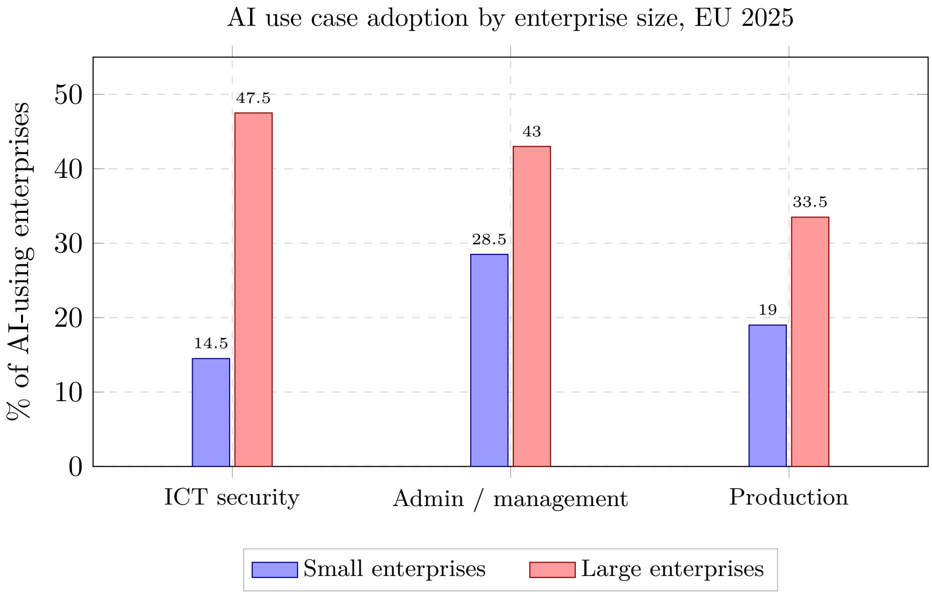


Figure 3.4. Share of AI-using enterprises deploying AI within each functional area, broken down by firm size. Across all three use cases large enterprises lead small enterprises by a substantial margin (+33 p.p. in ICT security, +14.5 p.p. in admin/management, +14.5 p.p. in production processes), reflecting both greater investment capacity and higher process complexity at scale. Source: Eurostat [28].

infrastructure (for RPA and analytics) or might not be immediately necessary for a small business (perhaps not every small business needs a customer service chatbot, whereas a large company with huge inquiry volumes might). Similarly, small firms were behind in data analytics using AI (about 10 percentage points lower adoption) and in deploying AI-powered chatbots or virtual agents (7 points lower)[38]. Robotic process automation (RPA) – which involves automating back-office workflows – showed a 16.7 percentage point lower adoption in small firms compared to large[38]. The pattern reinforces that SMEs pick and choose AI uses that align with their immediate needs and capabilities – typically customer acquisition and service improvement – and hold off on those that require re-engineering internal processes or heavy data lifting.

This suggests that small businesses tend to adopt AI at the surface level, leveraging ready-made solutions, whereas large firms more often undertake deeper integration and internal capacity-building. They were likely using AI tools out-of-the-box or as a minor part of their operations. Another aspect of "how" SMEs adopt AI is the depth of investment they make. Many small firms use AI in a lightweight manner – e.g., using a pre-built AI feature in software, but not investing in custom AI development or extensive training. By contrast, only about 40% of large firms using AI reported making no such investments – implying the majority (60%) of large companies did invest in dedicated AI talent, infrastructure or training[38]. A survey found that about 50% of small firms using AI reported making no additional investments in AI (such as training staff, hiring AI specialists, or purchasing new software/hardware)[38]. Indeed, a figure in that source

(Figure 4) shows large firms outpacing small firms in activities like training staff to use AI, hiring AI-trained staff, and changing data management practices to accommodate AI[38]. SMEs, often constrained by budgets, might skip those steps and use AI in a plug-and-play fashion – which can be efficient in the short term, but might limit the transformative impact of AI on the firm.

3.3.4 Barriers and Attitudes Affecting Adoption

Table 3.1. Main barriers to AI adoption reported by SMEs. Figures sourced directly from the 2025 enterprise survey [28] and recent reports on SME workforce skills [39] and on D4SME survey results [36], plus a systematic review of 27 barriers [29]. Categories partially overlap.

Barrier category	**Reported share / status**	**Source type**	**Reference**
Lack of relevant AI expertise (within firm)	70.9% of SMEs that considered AI but rejected it	Eurostat 2025	[28]
Skill / talent gap for generative AI	50%+ of non-GenAI SMEs cite skills as main reason	OECD survey	[36, 39]
Lack of clarity about legal consequences	52.5%	Eurostat 2025	[28]
Data protection / privacy concerns	48.8%	Eurostat 2025	[28]
Lack of time for staff training	43%	OECD survey	[39]
Lack of skilled talent broadly	27%	OECD survey	[39]
Lack of knowledge & high costs & inadequate IT infrastructure	Top-3 most-cited barriers (out of 27 identified)	Systematic review	[29]
Belief "AI not useful for our enterprise"	20.7%	Eurostat 2025	[28]

Understanding the state of AI in SMEs also requires looking at the non-adopters – why many SMEs still do not use AI. The top barriers are quite aligned with the structural challenges discussed above and can be summarized as lack of appropriate skills/knowledge, concerns about costs and return on investment, and uncertainty or fear regarding AI (including legal and data concerns). Surveys of enterprises consistently highlight a set of perceived barriers that hold back small firms from implementing AI.

SMEs are wary of the regulatory and liability aspects of AI , for instance, questions about who is responsible if an AI makes a faulty decision, or concerns about complying with data protection laws (GDPR in Europe) when using AI that processes customer data. In plain terms, SMEs felt they did not have employees with the necessary know-how (or access to consultants) to effectively implement and manage AI systems. The second most common reason was lack of clarity about legal consequences (mentioned by 52.5% of firms)[28]. This is understandable: hiring AI specialists is expensive and competitive, and smaller firms often rely on generalist IT support that may not cover cutting-edge AI skills. A 2025 enterprise survey asked enterprises that had considered using AI but ultimately decided against it about their reasons. The most common reason (cited by 70.9% of such enterprises) was a lack of relevant expertise to use AI[28]. It is worth noting only ~20.7% said a reason was that AI "was not useful for our enterprise"[28]. This indicates that the vast majority of SMEs do see potential usefulness in AI for their business – the issue is they feel impeded by lack of skills or external constraints, rather than thinking "AI has nothing to offer us." This is an important point: it's less about convincing SMEs that AI is useful (most already believe it could be) and more about enabling them to overcome practical barriers to adoption. Following closely, about 48.8% cited concerns regarding data protection and privacy specifically[28]. Clearly, uncertainty in the legal/regulatory environment and fear of violating privacy rules create hesitation for many SMEs.

Academic and industry studies echo these findings. In surveys, many SMEs express worry about the initial investment cost and uncertain return from AI projects, a calculation that larger firms with more slack can more easily take a chance on. Cost is a perennial concern: implementing AI can require investing in software, possibly new hardware or cloud services, and training or hiring staff, all of which strain an SME's budget. A comprehensive literature review published in 2024 identified 27 distinct challenges that SMEs perceive in AI adoption, with the most commonly cited being lack of knowledge, high costs, and inadequate infrastructure[33]. Lack of knowledge refers not only to technical AI expertise but also knowing where AI could fit in the business – many SME owners simply aren't sure what AI use case would provide ROI for them. This is often a misperception, but a prevalent one in certain segments. In a US survey, among businesses with fewer than 5 employees not planning to adopt AI, a striking 82% gave as the primary reason a belief that AI is not relevant or useful to them [38]. Only about 6–7% of those micro-businesses cited lack of knowledge or privacy concerns as their top reason; the vast majority simply felt that "AI isn't for us". For instance, some SMEs – especially the very smallest – have a mindset that AI is not relevant to their business. Inadequate infrastructure, as discussed earlier, is about not having the IT environment to plug AI into – whether that is data infrastructure or compute power. As firm size increases, that attitude becomes less common – bigger SMEs more readily acknowledge AI's relevance, instead their worries shift to practical issues like knowledge and cost[38]. These are the "hard" barriers; there are also "soft" barriers such as organizational culture and awareness. This highlights an attitudinal barrier at the very small end of the SME spectrum: a bakery with 3 employees or a local plumbing service might sincerely think AI has no role in their tiny operation, whereas a 50-employee manufacturing supplier might desperately want AI for quality control but can't afford it or find the talent. Therefore, the psychological and informational barriers vary with size and sector: outreach and education may be needed for the smallest firms to open their eyes to AI opportunities, whereas tangible resources and expertise are needed for the medium-sized firms that are already keen.

Building on this evidence, Chapter 3 will revisit these barriers in a more structured way, clarifying how they interact and how they shape the feasibility of adoption across different SME contexts. It will also briefly reflect on how some historical constraints have started to ease in recent years as AI capabilities have become cheaper and more accessible through widely available tools and models (including open approaches), even though skills, data readiness, and regulatory uncertainty remain central challenges.

At the same time, the majority of SMEs have yet to adopt AI, and they face a set of interlocking barriers – primarily skill and knowledge gaps, costs, and uncertainty – that need to be addressed to unlock wider adoption. The experience of those adopters seems generally positive, with many finding multiple uses for AI within their businesses. In terms of the current state, AI adoption in SMEs, while still trailing large companies, is on a clear upward trend. There are success stories of small firms using AI to personalize marketing at scale, to forecast demand better, or to automate tedious tasks, yielding improvements in efficiency or sales. A growing minority of SMEs have begun to integrate AI tools, mostly in areas like marketing, customer service, and data analysis. Policymakers in Europe and elsewhere are aware of this: the EU's Digital Decade goals explicitly aim for 75% of companies to use AI, cloud, or big data by 2030[28], which implies a major push to bring SMEs on board (since most large firms already do). Achieving such targets will require lowering the barriers for SMEs, for example through training programs, clear legal frameworks, and affordable AI-as-a-service offerings.

The next chapter of this thesis will examine what can be done to help more SMEs cross the threshold into AI usage. There is a gap between the enthusiastic headlines about AI and the on-the-ground reality for many small businesses; yet there is also a cohort of SMEs that have embraced AI and are proving its value in a small-business context. In conclusion, the state of AI adoption in SMEs as of 2025 can be characterized as early-stage but accelerating. But before that, in the final section of this chapter, the final section considers why AI adoption by SMEs matters – what are the implications for SME competitiveness and for the economy, and how does this link to the broader themes of the thesis.

3.4 Implications for Competitiveness of SMEs

The adoption (or non-adoption) of AI by SMEs is not just a technical or internal matter – it has broad implications for the competitiveness of individual firms and for economic development as a whole. This section reflects on those implications and connects them to the objectives of this thesis. This dynamic will influence the competitive landscape in many industries and could widen or narrow the "digital divide" in the business sector, depending on how it unfolds[37]. The central argument is that AI represents opportunity and threat for SMEs: those that harness it effectively can potentially punch above their weight, increasing productivity and innovation, whereas those that fall behind risk exacerbating the performance gap between themselves and larger or more tech-savvy competitors.

AI can enhance SME competitiveness in several ways. This is particularly important for SMEs, which as noted above tend to have lower labour productivity than large firms[24]. By deploying AI (for example, an AI system that automates scheduling or quality inspection), an SME can get more output from the same input, closing some of the productivity gap. From a firm-level perspective, competitiveness is about the ability to deliver better products and services, more efficiently, and to adapt to market changes. AI can also enable new capabilities that small firms previously lacked – for instance, advanced predictive analytics for demand forecasting or personalized recommendations for customers were once the preserve of big companies with data scientists, but AI-as-a-service now allows SMEs to access such capabilities, potentially improving their decision-making and customer experience. For one, AI tools can increase productivity by automating routine tasks and optimizing processes. A single customer support agent in an SME can handle more queries with the help of an AI chatbot triaging questions; a small marketing team can achieve broader reach with AI-driven ad targeting; a lone inventory manager can maintain leaner stocks with AI prediction of sales. In effect, AI can act as a force multiplier for a small firm's limited resources. If well-used, AI allows SMEs to scale their effectiveness without proportionally scaling headcount or costs, which is a huge competitive boon in tight markets.

Also, AI can spur innovation in business models and offerings [40]. Crucially, a majority (57%) of those businesses saw more opportunities than risks in using generative AI[39]. SMEs are often very innovative in niches and can be agile in trying new ideas. This optimism is likely because generative AI lowered barriers – it often requires only basic prompting skills and is available at low or no cost for trial, so even firms with very limited tech infrastructure could test it. With AI, they may develop new services (for example, an SME consultancy offering AI-augmented analysis to clients, or a manufacturing SME moving into smart customised production using AI). Generative AI, which saw a breakthrough in 2023, is a case in point: within a year of its public release, nearly 1 in 5 businesses (including many SMEs) were already experimenting with generative AI tools[39]. It also suggests that some previous disadvantages (like not having graphic designers on staff – now an SME owner can use AI image generation for graphics) can be mitigated by AI, if adopted. Small companies have been observed using generative AI to create marketing content, design prototypes, or assist in coding, which can speed up development cycles and reduce costs. This democratization of sophisticated capabilities means SMEs that are willing to experiment can gain innovative edges quickly.

There is a concern that a digital divide within the business community will lead to widening inequality in firm performance[35]. It has been warned that "too many SMEs have yet to 'go digital' and risk being left further behind with the launch and widespread adoption of generative AI models"[39]. Early-adopting firms (often larger or more resource-rich) can improve efficiency faster, cut costs, and capture market share, while late adopters stagnate. In other words, each new wave of technology (first basic ICT, then e-commerce, now AI) has the potential to increase the gap between the frontier firms and the laggards. However, the flip side is that if SMEs do not or cannot adopt AI, they risk losing competitiveness relative to those who do (including large enterprises and more digitally advanced peers). SMEs collectively account for a huge portion of employment and value-add – if they underperform, it drags down overall productivity and innovation in the economy[24]. This is why institutions like the European Commission have made diffusion of digital tech and AI among SMEs a policy priority, tying it to Europe's competitiveness in the global arena[28]. The EU's goal of 75% of companies using AI or big data by 2030 is fundamentally about not letting a large base of SMEs fall behind in the next decade's innovations[28]. On a larger scale, entire economies could suffer if their SME sectors remain digitally backward. If a local small retailer never adopted e-commerce, it likely struggled when consumers moved online; similarly, if now it doesn't leverage AI for, say, demand forecasting or personalised marketing, it could struggle to compete with a rival that does and so can manage inventory better and attract customers more effectively.

A study on SMEs during crises found that those who had embraced digital tools felt better equipped to weather disruptions – in one UK survey, 74% of SMEs said technology adoption left them better prepared for future crises[35]. The past few years, marked by the COVID-19 pandemic and other shocks, showed that digitalized businesses were more resilient – they could pivot to online sales, remote work, etc., more easily. Another important implication is on the resilience and adaptability of SMEs. SMEs that integrate AI into their operations (say, an AI system monitoring market trends or a predictive maintenance AI that reduces downtime) will likely be more agile and capable of absorbing shocks than those running entirely by manual processes and gut instinct. AI can further enhance resilience by, for example, helping firms respond to disruptions (AI-driven scenario planning, intelligent supply chain adjustments). Conversely, SMEs lacking digital maturity can find themselves more vulnerable when a crisis hits, widening the survival gap.

AI adoption will change the skill needs in SMEs [39]. It's also worth considering competitiveness in terms of labor and skills. Firms that adopt AI may need fewer workers for certain routine tasks but more for higher-skilled roles (like data analysis, AI oversight, or tasks that AI cannot do). On one hand, this might disadvantage SMEs in the labor market competition for AI talent (big firms can pay more). There is a risk that SMEs could face a skill crunch – they already report lack of skilled workers as a barrier to digital adoption [39]. There is an implication for workforce competitiveness: SMEs will need to invest in training to use AI effectively. Currently, many SMEs rely on informal learning (on-the-job, peer learning) for digital skills[39]. On the other hand, if SMEs do manage to upskill their existing workforce, they could become quite attractive places to work – offering employees the chance to work with cutting-edge tech in a small-team environment. Overcoming this will be key. Conversely, SMEs that cannot navigate the retraining of staff or integration of AI and human workflows might see workforce issues (either job losses or inability to utilize AI well, leading to no productivity gains). As AI reshapes roles, the most competitive SMEs may be those that effectively combine human talent with AI – using AI to augment their staff's capabilities rather than replace them. The evidence shows that lack of time for training is cited by 43% of SMEs as a bottleneck to digitalisation, and lack of skilled talent by 27%[39].

As a result, ensuring that SMEs can overcome the hurdles and leverage AI is of critical importance. There is evidence of win-win scenarios where improving SME productivity also benefits larger firms through supply chain linkages and overall economic vibrancy[24]. But done wrong (or not at all), AI could exacerbate existing divides, leaving many small firms struggling and economies missing out on a major source of potential growth. Ultimately, empowering SMEs with AI is not just about technology – it's about equipping the most numerous class of businesses with the means to thrive in the future economy, thereby fostering broader competitiveness and innovation. In summary, the implications of SME AI adoption for competitiveness are profound. AI could help close the SME productivity gap and enable small firms to remain viable and dynamic in a digital economy – contributing to inclusive growth and innovation diffusion.

The chapter has outlined who SMEs are and why they matter (2.1), how they operate and where their key processes lie (2.2), what structural issues hamper their digital progress (2.3), the current extent and nature of their AI adoption (2.4), and why it is crucial to bridge the remaining gap (2.5). The next chapters will first systematise the evidence from the literature by organising the recurring challenges and needs of SMEs into a clear analytical picture, so that the problem space is defined in a consistent and comparable way across firms. This concludes the chapter's examination of SMEs and AI adoption. Armed with this understanding, the remainder of the thesis moves from describing the problem to structuring and addressing it. Building on that, subsequent sections will then discuss how SMEs can progress from low digital maturity and fragmented processes towards more effective use of AI, focusing on the conditions that make adoption feasible in practice and the types of support or enablers that reduce typical barriers (e.g., skills, data readiness, integration, and uncertainty). Finally, the later chapters will consolidate these insights into a coherent set of implications and guidance, connecting the earlier diagnosis to the broader objectives of the thesis and outlining what this means for SME competitiveness and the adoption journey over time.

Chapter 4

Adoption Barriers and How They Have Lowered

The present chapter documents how structural obstacles to AI have been progressively eroded for small and medium enterprises (SMEs) between 2020 and 2024, and identifies the barriers that persist after that process. The analysis weaves together three empirical strands: the pre-2020 baseline of barriers documented in qualitative work on Industry 4.0 adoption among multinationals and SMEs [41], a 22-expert qualitative pilot across the Czech Republic and Austria [42], and a 2025 global survey of 1,993 firms across 105 countries [43]. Five ecosystem shifts are then mapped onto these barriers: the LLM-as-a-service price collapse [44–46], the open-weight and edge-inference movement that ended structural vendor lock-in [17, 47, 48], no-code platforms enabling AI integration by domain experts [49–52], and the vertical SaaS layer that bypassed integration for many use cases [53–58]. The chapter closes on the residual organisational barriers, namely skills, data governance, and pilot-to-production scaling, which are not resolved by technology alone [59–62]. The barrier reductions established here supply the precondition for the problem taxonomy of Chapter 5 and the solution-and-build-versus-buy mapping of Chapter 6. Throughout this chapter, "AI" denotes chiefly large language models (LLMs) and the broader category of generative AI, as distinct from the classical machine learning (ML) that preceded them; the term is used in this narrower sense except where the wider field is explicitly intended.

4.1 Historical AI Adoption Barriers for SMEs (Pre-2020 Baseline)

Before 2020, the baseline consisted of mutually reinforcing barriers that operated in concert to keep SMEs away from AI. The literature converges on a set of recurring barriers - cost, skills, infrastructure, and integration complexity - with most sources referring to the same or closely overlapping sets of obstacles [41, 42]. Each barrier was separately identified as an isolated obstacle, and each targeted the resource asymmetries between SMEs and large firms. Cumulatively, however, they amounted to an effective barrier to entry against a technology that was, in principle, already available.

Before turning to these barriers, the population scope of the problem deserves clarification. As defined by the European Union, small and medium-sized enterprises are any firm that is under 250 employees and which has an annual turnover that is less than €50 million [42]. In 2022, there were 24.3 million SMEs in the EU-27, accounting for 99.8% of businesses in the non-financial business sector. Similar concentrations of SMEs in non-financial business sectors recur across other OECD economies as well, with small and medium-sized firms making up the majority of private employment[61]. Without these enabling shifts, AI would have remained the preserve of firms large enough to fund in-house research and major capital projects. Following the OECD, an AI system is understood here as a machine-based system that, given explicit or implicit objectives, generates outputs such as predictions, recommendations, or decisions capable of shaping real or virtual environments [61]. Therefore, the discussion is not simply about what AI should do with a small subset of the economy; rather, it concerns whether AI will become a general-purpose technology - one whose use is pervasive across sectors, that improves continually over time, and that spawns complementary innovation in the activities that adopt it [63] - or whether it will remain a frontier technology for incumbents alone [36].

High cost was one of the top challenges for advanced-technology implementation in SMEs [41]. Due to their limited capital, the small business was unable to sustain the initial capital expenditure required to create, purchase, and implement AI systems. In addition, there would be an incurred cost for ongoing maintenance and upgrades to ensure the AI system operated correctly. In fact, SMEs often struggled to get the funding required for innovation, such as R&D, adopting new technologies, and hiring qualified talent.

Cost was not incidental to the era; rather, it was intrinsic to the way AI was developed and disseminated. Building on prior work, a qualitative pilot identified twenty-one critical digitalisation and innovation barriers total, with the highest three by scores being software/hardware costs, the complexity of technology selection and absence of government incentives, R&D support, policies, and standards [42]. The same constraints recur. Costs accumulated and AI required not only deployment but also retraining and regular updates [41, 42].

4.1.1 The Skills Gap and the Scarcity of Specialised Talent

Specialised talent is the second building block. High-level skills are required for AI and advanced digital systems to be deployed successfully, such as in software engineering, data analysis or process integration, and there is a shortage of this kind of skills available within the available labour supply of SMEs, creating a persistent talent asymmetry[41]. SMEs had problems hiring and retaining skilled staff and the main reason for this was limited resources which were not sufficient to compete with the offers of larger companies. The economics were straightforward: AI talent commanded premium compensation while SMEs operated on slim margins, so the two rarely aligned.

The same trend was revealed in twenty-two qualitative interviews with industry experts in the Czech Republic and Austria, where the four canonical barriers were identified as: lack of trust, lack of knowledge, infrastructural problem, and lack of skilled workers [42]. The knowledge gap occurred on two levels. SME owners and managers were unsure to an extent that they were clear about AI's benefits, or how AI could be applied in their sector, leading to an investment dilemma in terms of confidence. Under the implementation gap also lay a shortage of personnel that could construct, run, and utilise AI, meaning that despite some being clear as to their strategic goals there was still a shortage of human resource to act on them [41].

Finally, the third aspect was the constraint of infrastructural limitations. The dependence of AI applications on huge datasets triggered a debate over data sovereignty and compliance with the EU General Data Protection Regulation [41]. These SMEs had the onerous task of building solid data governance practices - the policies, roles, and controls that govern how data is collected, stored, accessed, and used - to ensure data is handled in an ethical and legal manner while meeting a tight framework of standards and policies. This is an extremely time-consuming process for organizations of such magnitude. In comparison, the infrastructure problems were described [42] simply as slow computing power, outdatedness and inadequate technical help.

This gap remains today. Figure 4.1 shows that, in a survey conducted 25 June 2025-29 July 2025 on 1,993 respondents in 105 countries, 88% say they are using AI at some level of their organization across any business function, up 10% from the prior year's 78% [43]. This scaling gap is still by size, with almost 50% of the companies with revenues in excess of US$5 billion in the scaling phase, compared to 29% of those with less than $100 million. A third of respondents are scaling use of AI but the majority were still piloting or experimenting. Where AI is being used only 39 percent say it is impacting EBIT and only a small portion of these that AI is contributing more than five percent. The barriers have not disappeared but have changed form; the pre-2020 baseline is therefore not merely a historical point but a benchmark against which the following chapters are measured.

4.2 The LLM-as-a-Service Revolution and API Economics

The first change in the structure of the post-2020 AI ecosystem, the one for which the empirical record is most granular, is economic. A large language model (LLM) is a generative model trained on very large text corpora to predict the next token in a sequence, delivered to applications through three distinct tiers: the trained model, the surrounding infrastructure, and the end-user application [45]. LLMs are thus one class of generative AI, which is itself a subset of artificial intelligence (LLM $\subset$ generative AI $\subset$ AI); they are distinguishable from the classical, task-specific machine learning that preceded them in both scale and capability, as the latter is trained on narrow labelled datasets for a single predictive task rather than on broad corpora for open-ended generation. In terms of the rate of decline in the per unit price of running such a model, the years 2021 to 2024 saw a price collapse of roughly a factor of ten per year, constituting a thousand-fold reduction in three years. However, the price collapse was not uniform; the decline was concentrated in the non-frontier tiers, while the most capable frontier models remained comparatively expensive [44, 45]. As a result of this change, the cumulative price reduction crossed the investment threshold that had previously kept SMEs away from large-scale AI implementation.

An empirical study of LLM inference costs [44] compiled the pricing data that was

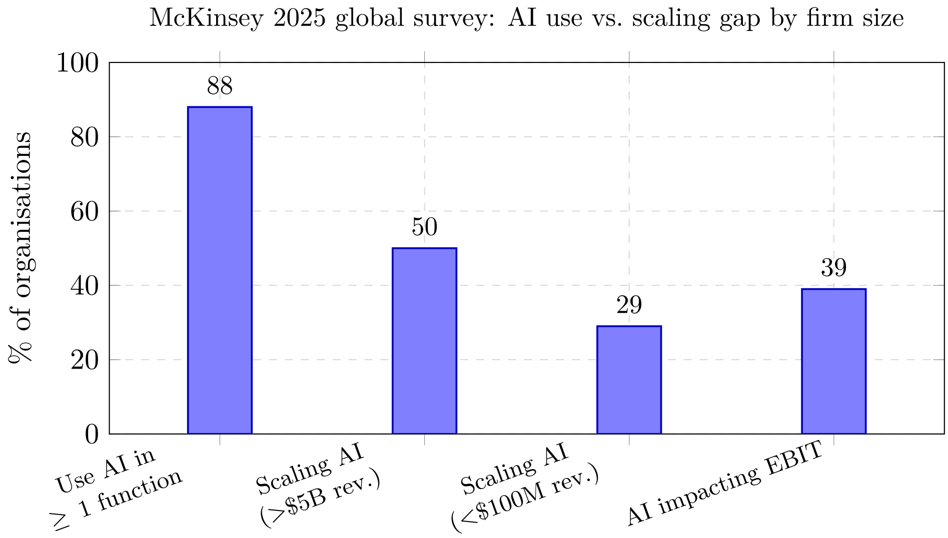


Figure 4.1. The persistence of asymmetry: 88% of organisations report using AI in at least one business function, yet only one in three is scaling, with a 21 percentage-point gap between firms above $5B in revenue and those below $100M. Only 39% of respondents perceive an EBIT impact, indicating that adoption breadth has outpaced operational depth. Source: McKinsey *State of AI* 2025 global survey, 1,993 respondents across 105 countries [43].

available at the time of writing for OpenAI, Anthropic, and Together.ai. The study examines the inference costs of a representative set of LLMs - OpenAI's GPT-3, GPT-3.5 Turbo, GPT-4, and GPT-4o, Anthropic's Claude models, and LLaMA 3.2. As depicted in Figure 4.2, from January 2022 to mid-2024, the cost to process a million tokens plummeted from $60.00 for GPT-3 to $0.06 for LLaMA 3.2 3B, a thousand-fold reduction in three years. Specifically: $60.00 in 2021, $20.00 in 2022, $2.00 in 2023, and $0.06 in 2024. In 2023, Anthropic released Claude 1 with a rate of $55.00 for a million tokens, mirroring the introduction rate for GPT-4 at $10.00. The starting price in 2024 was between $0.06 and $2.00 for a million tokens [44].

Three inflection points in price are suggested: $2.00, $0.10, and $0.06 for each million tokens [44]. These price points correspond with categories that arise at those levels, ranging from document automation and chatbots to multimodal analytics platforms. What is significant about these price points is that they are a function of which products can be economically supported by the price. At $0.06 per million tokens, a chatbot that processes 500 tokens per interaction will cost 0.003 cents per interaction. In 2021, the same operation at GPT-3 prices would have cost roughly $0.03.

These structural shifts in pricing are connected to the very design of the pricing models themselves. From 2021 to 2024, the industry transitioned from set-rate pricing to a matrix of pricing which includes model size, task complexity and tier of service [44]. That difference is important for operations, too, as it represents a shift from AI as a capital expense, buying, training and deploying a model, to an operating expense,

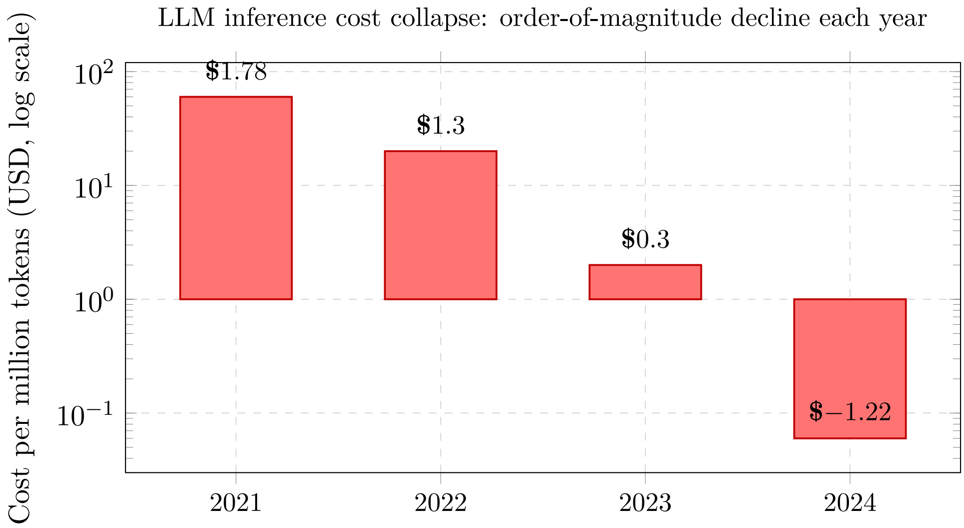


Figure 4.2. The price collapse of LLM inference, 2021–2024. Frontier-tier models (GPT-3 in 2021) priced at \$60.00 per million tokens dropped roughly a factor of ten each year, reaching \$0.06 (LLaMA 3.2 3B entry-level) in 2024 - a 1,000-fold reduction over three years. The log-scale axis reveals the geometric, not linear, nature of the decline. Source: Sysoiev (2025) [44].

paying per token. For small-to-medium enterprises, the distinction is a meaningful one. It represents the difference between securing capital-expenditure approval and adding a recurring operating-expense subscription - a shift from capital expenditure to operating expenditure [18].

4.2.1 Startup Formation, Capital Flows, and the Demand Side

But that is only the supply side; the demand response tells the complementary story. The same study correlates pricing information with startup-formation and capitalization information from sources such as Crunchbase, Quid, and FDI Intelligence[44]. The number of AI start-ups has grown by a factor of 1.7 between 2021 and 2024, with 1,200 new AI start-ups founded in 2021, 1,289 in 2022, 1,890 in 2023, and 2,049 in 2024. The total amount of capital invested has gone from \$121 billion to \$151 billion between 2021 and 2024, with a 44.5% year-on-year increase in 2024. The fact that the numbers decreased in 2022 and 2023 (to \$88b and \$104b respectively) and then increased again in 2024 to reach \$151b is a sign that investors were willing to commit capital to the AI-as-a-service business model. The results, summarised in Figure 4.3, show that price decrease has stimulated the creation of start-ups, and a similar increase has also happened for venture investment volume [44].

Generative AI has been characterised as a socio-technical object composed of three layers: the model, the system and the application [45]. This framework is helpful in understanding why the API distribution model is not simply a model of distribution. The model layer is formed of trained weights and a probability generation mechanism. The

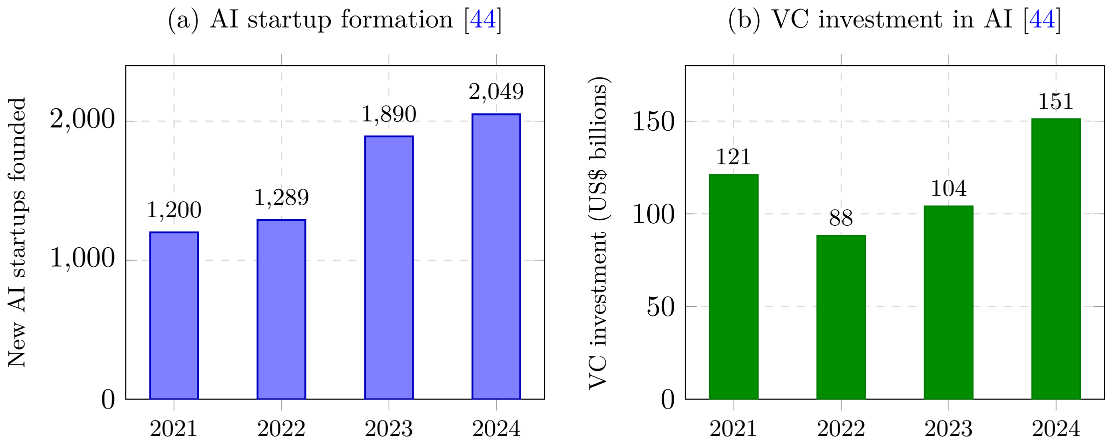


Figure 4.3. Demand response to the price collapse. (a) AI startup formation grew by a factor of 1.7× between 2021 and 2024, from 1,200 to 2,049 newly founded firms per year. (b) Venture-capital investment dipped during the macroeconomic correction of 2022–2023 before rebounding to $151 billion in 2024 (a 44.5% year-on-year increase), signalling investor confidence in the AI-as-a-service business model. Source: Sysoiev (2025) [44].

system layer adds data processing, retrieval and orchestration. For SMEs, the application layer is the point of contact with the actual AI: from customer service, content, code generation and support in decision making [45]. Each layer of the model can be priced and used by the market separately, which is what API economics allows.

4.2.2 API Demand and Industry Forecasts

In combination, the decline in costs on the supply side and the rapid increase in startups on the demand side have yielded a third major discovery: the nature of API usage is shifting toward AI. In a poll of 459 technology service providers that ran in October, November, and December of 2023, it was concluded that 30% of the growth in API demand by the year 2026 would come from tools that use artificial intelligence and large language models [46]. The growth mentioned here is not of some generalized nature. It means that demand for API services is now, to some extent, being driven by AI as opposed to other web and mobile use cases. While the poll measures the composition of API demand rather than the experience of buyers, this shift plausibly implies, for SMEs that are consumers rather than creators of APIs, a more competitive market with a wider range of options - though the claim that switching friction falls as a result is an inference rather than a measured finding.

None of the dynamics traced above removes every adoption hurdle: the preceding discussion addresses only the financial barrier. The remaining sections of the chapter turn from price to the other three pillars - vendor lock-in, technical integration, and the lingering organisational challenges that even free AI will not resolve.

4.3 Open Source Models and the End of Vendor Lock-In

Cheaper inference does not, on its own, confer independence. An SME which pays $0.06/1M tokens to one single cloud is still strategically at risk: if that cloud changes pricing, or policies, or roadmap, the effect propagates directly into the SME's operating model. This dynamic is what the second major shift of the post-2020 period directly mitigates. Open-weight models, locally-runnable inference, and the broader open-source movement all contribute to providing an alternative to deploying closed-source. That transition is far from complete, yet the path is now both technically and economically viable.

4.3.1 The Concentration Baseline: Cloud Platform Power

This is not merely an abstract issue regarding vendor lock-in. As characterised in *Platform Power in AI: The Evolution of Cloud Infrastructures in the Political Economy of Artificial Intelligence* (Internet Policy Review), it is a concentration, a measurable phenomenon. AWS, Microsoft Azure, and Google Cloud, respectively, make up 68 percent of the global cloud-computing market, a level of platform concentration with direct implications for switching costs, according to the newest market data published in 2024 [17]. All three services provide a three-layer cloud stack, including Infrastructure-as-a-Service to compute, store, and network; Platform-as-a-Service to develop, train models, and create software development kits; and Software-as-a-Service for proprietary, end-to-end AI offerings. Each layer increases the footprint of the platform and generates switch costs that are self-reinforcing.

Figure 4.4 highlights three dimensions of infrastructural power: vertical integration, complementary innovation, and the power of abstraction [17]. Vertical integration may introduce path-dependencies that result in lock-in. Technical boundary resources, such as APIs, SDKs, developer documentation, and machine-learning libraries, force developers to follow the ecosystem conventions of a particular platform broker, which in turn grants these platforms the power to charge a take-rate on every transaction. The compute layer further amplifies this asymmetry. As further noted in the same work, concentrated ownership of graphics processing units by the major platform players catalyses a compute divide that gives these players, together with their academic collaborators, a distinct competitive advantage. On an even larger scale, Microsoft's partnership with OpenAI, as well as Google's purchase of DeepMind, have seen both foundation models brought in-house on top of platform infrastructures that already dominate the compute layer, in the form of DALL · E and PaLM respectively.

For the SME, this means that its production environment can end up highly dependent on the AI infrastructure provided by one particular cloud service provider. This dependency will not be equal across the board; that is, an SME's IaaS-level reliance will be easier to relocate than its dependence at the PaaS level, which in turn is likely easier to relocate than a reliance on SaaS-based, provider-specific AI capabilities. So the choice of how to adopt depends in part on the amount of risk the SME is willing to absorb at various levels. Whether a given SME should build, assemble, or buy a capability, and which SME categories fall on each side, is the subject of the build-versus-buy analysis in

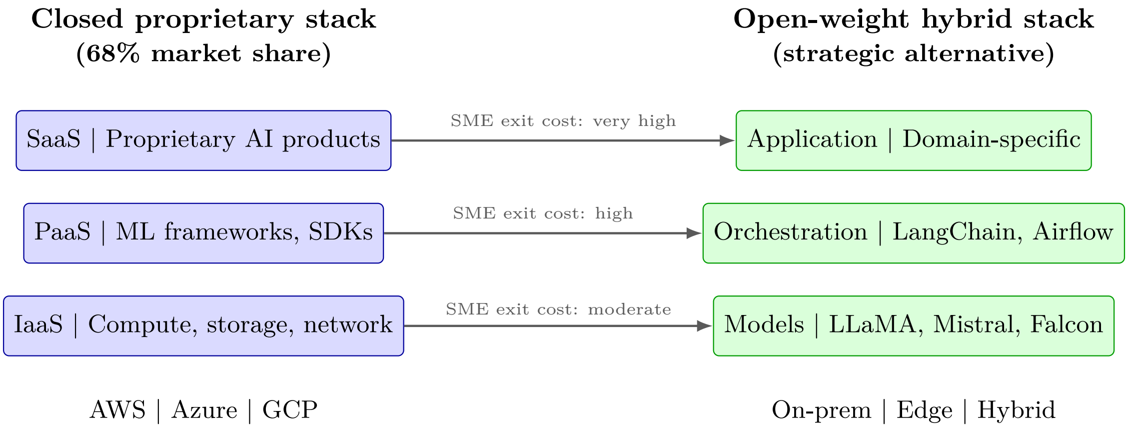


Figure 4.4. Cloud platform concentration vs. open-weight alternative. AWS, Azure, and GCP control 68% of the global cloud market [17] and offer a three-layer stack (IaaS/-PaaS/SaaS) whose switching costs rise as the SME moves up the stack. Open-weight models (LLaMA, Mistral) deployable on-premise, at the edge, or in hybrid configurations restore portability without forcing a complete abandonment of cloud services. Adapted from Luitse [17], with deployment routes from Zhang et al. [47] and Pereira et al. [48].

Chapter 6 and of the technology-selection phase in Chapter 7; this chapter establishes only the enabling conditions that make the "buy" and "assemble" options viable.

4.3.2 Open-Weight Models as Strategic Alternative

The technological counter-response to platform power is the proliferation of open-weight models. open-weight means that the trained parameters of a model can be freely downloaded by an individual to self-host them. This is a more limited sense than fully open-source, which would also require that the training data and code be made available. This means that a model can be open-weight (the weights are public) and not open-source in the more strict sense. The value of open-weight models is explicitly called out, for example that models like LLaMA and Mistral provide a different set of benefits than the commercial providers: transparency, customisability, collaboration, and cost [48]. Prompts made to ChatGPT, Mistral and LLaMA are examined to demonstrate how using established benchmarks allows direct comparison of these models on a range of shared tasks. Generative AI, particularly large language models, has the potential to be leveraged on software-engineering tasks including automating coding, design and documentation. However, the potential of AI software engineering also presents reliability considerations that should be explored in terms of their effectiveness and consequences.

The strategic point is that the choice between proprietary and open-weight is no longer binary. SMEs can route latency-sensitive or privacy-sensitive workloads to an on-premises instance of an open-weight model, while keeping more general-purpose workloads on the API. A hybrid model dilutes the risk and prevents any single vendor from accumulating excessive leverage over pricing.

4.3.3 Local Inference as Practical Deployment Path

Another lingering concern is whether open-weight models can truly scale and operate at latencies that make them viable for production environments. This concern is addressed in the work *EdgeShard*. LLMs are commonly deployed on cloud servers, which introduces high latencies, expensive bandwidths, and privacy issues [17, 47]. Because edge devices are located near the data source itself, edge computing is seen as a possible solution for these problems. EdgeShard is presented as the pioneering method that utilizes LLMs in an integrated, collaborative edge-computing platform. In a collaborative edge-computing framework, edge devices and cloud servers coordinate to jointly provide an efficient LLM inference service without incurring any loss in accuracy. When evaluated using Llama-2 serial models in an operational testbed, EdgeShard can offer up to a 50 percent reduction in latency, and boost the overall throughput by 2×, compared with current state-of-the-art methods, as shown through comprehensive experimentation [47].

The significance is that it eliminates the necessity to choose between dependence on the cloud and accuracy. SMEs concerned about data sovereignty, latency, or vendor lock-in have a documented path to local and hybrid deployment of open-weight models. Operational challenges persist: edge system management, prediction arrangement, disaster recovery, and model updates all need engineering work, however the documented functionality is in line with entirely cloud based ones [47]. Along with the expense patterns discussed in the preceding section, the open-weight option makes vendor dependence a decision as opposed to an inherent necessity.

4.4 No-Code Platforms and the Democratisation of Integration

If affordability and vendor lock-in are pillars one and two in the legacy exclusion of SMEs, then the integration complexity constitutes the third. Even with cheap inference and portable models, organizations that depend on bespoke engineering to integrate AI into their data flows, business processes, and decision-making pipelines will encounter a constraint that scales inversely with the engineering resources available. This is the exact problem the no-code and low-code ecosystems that have emerged since 2020 are designed to overcome. No-code platforms let users build applications entirely through visual, configuration-driven interfaces without writing code, whereas low-code platforms reduce, but do not eliminate, the hand-written code required. Today, configuring an AI workflow is now closer to the competence of a domain expert than to that of an engineer.

4.4.1 Orchestration Frameworks: Adaptive AI Pipelines

Complementing this transition to a code-free approach is a corresponding development in terms of the orchestration layer. MLOps (machine-learning operations) denotes the set of practices and tooling that automate the deployment, monitoring, and lifecycle management of models once they leave development and enter production. A comprehensive analysis of MLOps platforms and tools, recently published in *Artificial Intelligence Review*, links several open-source orchestration frameworks, including Apache Airflow, Kubeflow, and Prefect, with various operational responsibilities encountered after a model leaves the testing phase [51]. Tools are classified by the functions they perform, including designing pipelines, model training and evaluation, releasing, and tracking, and the advances made in these fields between 2020 and 2024 are found to have significantly reduced the amount of engineering time spent developing AI pipelines at scale.

From an operational perspective, this shift matters most to firms with limited human resources. Workflow orchestration tools enable the ordering of task executions based on factors such as dependencies, availability of resources, and past reliability, thereby minimizing delays and performance bottlenecks while allowing runtime modifications to workflows in response to the emergence of new services and data sources. The same report states that with well-developed orchestration systems, companies are able to configure their data processing in a custom way and adjust the scale of their resources as needed to fit in with the needs of both large corporations and smaller ones [51]. Furthermore, by automating the coordination of data pipelines and task activities, orchestration frameworks can reduce the reliance on manual input, the likelihood of errors and the elapsed time needed to transform a pipeline to accommodate changing requirements.

4.4.2 Embedded Productivity Tools: Microsoft Copilot in the SME Context

The third tool family embeds AI into existing productivity software through the no-code movement. The 2024 Work Trend Index anchors these trends empirically with thousands of data points. The 2024 study sampled 31,000 knowledge workers across 31 countries

[52]. It reports that 75 percent of knowledge workers are using AI on the job. It found that among those using AI at work 78 percent are employing their own AI tools (BYOAI). The survey further suggests that 79 percent of leaders surveyed feel the need to adopt AI in order to keep pace with their peers. The study reports that those classified as advanced AI adopters save in excess of half an hour per workday relative to non-adopters.

Copilot for Microsoft 365 showcases the embedded productivity design enabling this leap: a multimodal generative AI that takes text in, returns text and images, and employs a grounding pipeline leveraging Microsoft Graph to anchor answers in organizational documents, calendar entries, and email messages (if the user has access) [64]. It coordinates large language models across a single conversation.

Copilot for Microsoft 365 is designed to fit within the budget constraints of SME applications. It is offered as an annual add-on subscription for companies that already have Business Standard, Business Premium, E3, or E5 Microsoft 365 licenses [64], which means there is no need to maintain an additional AI framework separate from the company's current software environment. This enables SMEs to quickly adopt the service, bypassing the need to create new AI systems from the ground up or search for different vendors to work with.

Across these three layers a single pattern emerges: the integration boundary shifts closer to the domain expert, while the technical surface exposed to the user shrinks progressively. As observed of low code and no-code platforms, the challenge is less in the capacity to code and more in the capacity to govern, and the winning teams are those who can balance the speed of low-code technology with appropriate security and control regimes [50]. The bottleneck for SME adoption at this layer is organisational readiness, not engineering capacity.

4.5 Vertical SaaS with AI Embedded: The Zero-Integration Path

The final pathway considered here is also the most relevant for SMEs without any in-house engineering capacity: vertical software, or software-as-a-service, that offers artificial intelligence as a feature. Vertical SaaS is software designed for a specific industry or function (for example legal, retail, or manufacturing), as opposed to horizontal SaaS, which serves a generic need across all sectors; the vertical variant has its domain workflows and data models pre-loaded, so AI arrives ready-configured for the sector rather than as a general-purpose tool. In this case, the integration problem vanishes since AI capabilities are built into industry-specific software. The SME is not configuring AI but rather utilizing software-as-a-service that already comes with integrated artificial intelligence.

4.5.1 The Macroeconomic Value at Stake

The April 2026 report *Executive Perspectives on AI and the Future of Software* calculates the opportunity at the layer that is accessible to SMEs through vertical SaaS; the total addressable market for vertical and domain-specific agents, workflows, and systems of record is approximately \$1 trillion to \$2 trillion globally [54]. This layer of the market is already experiencing competitive disruption by new entrants. The software incumbents have earned \$3.5 billion in revenue from Vertical AI. In turn, new startup competitors in the same space have raised \$7.3 billion, which indicates that new entrants currently have about twice the share of revenue of the incumbents in the segment at this early stage. The report frames vertical and domain-specific software as the central competitive arena for application vendors. Regarding the readiness of SaaS incumbents to the displacement risks described above, 13% of SaaS incumbents are currently scaling AI capabilities, 42% are piloting AI capabilities, 34% are experimenting, and 11% are planning [54]. For existing SaaS vendors that can successfully integrate AI within their existing products, there is an opportunity to uplift cross-sell rates, which can increase average contract value to 1.5× to 2× (or higher) for SaaS vendors that offer an AI feature as part of the baseline subscription. From an operational perspective, there is an implication for SMEs. The vertical SaaS market is in the process of significant reorganization; the platforms that aggressively pursue embedding AI are likely to be the ones to have the highest capabilities embedded in their offerings by the next renewal cycle.

On a related note, predictions on the impact of embedded AI on cloud enterprise resource planning (ERP) applications show an impact on productivity. Embedded AI is forecast to enable financial close to occur 30% faster by 2028 in cloud ERP applications, according to a February 2026 press release [55]. According to the same forecast, 62% of cloud ERP spending in 2027 will be directed toward AI-enablement, compared to 14% in 2024.

4.5.2 Domain-Specific Verticals: Legal, CRM, and Manufacturing

The macroeconomic forecasts focus on specific industries. A 2021 *University of Pennsylvania Law Review* article discusses AI-enabled tools in civil procedure as well as the potential future of adversarialism [53]. The legal-tech vertical serves as an example of how AI is applied to a highly regulated field in which humans oversee, and the legal profession is accountable, the technology.

CRM provides a more recent example of vertical SaaS. Salesforce, as a dominant platform, creates a shared multi-tenant system, where AI capabilities like Einstein lead-scoring, case classification, and workflow automation are presented as adjustable products. Evidence from an evaluation of a secure multi-tenant AI system on Salesforce found double-digit accuracy gains compared to rule-based approaches when performing lead-scoring and workflow-execution tasks [56]. This demonstrates the potential benefits to a mature vertical SaaS that can offer AI capabilities. The practical implication for a small company is that an AI-enhanced CRM is consumed rather than constructed.

A third vertical, manufacturing, illustrates the same pattern. A paper on domain-specific MT for SMEs recounts how an SME with no in-house ML/AI capacity assembled a domain-specific machine translation system from about 5,000 PDFs of technical manuals [57]. A crucial part of the data-acquisition step was identification of sources, owners, rights of use, and means of access before training can even start. Information from different platforms and providers, including numerous outside suppliers and data lakes, was aggregated from many formats (some proprietary and some cloud based). And though such companies utilize the system, they remain cautious about translation quality, since domain-specific terminology often diverges from general language and may create reliability concerns in operational use. The vertical-SaaS route does not eliminate the data-preparation burden; it removes the engineering labour surrounding the data.

4.5.3 Residual Constraints and the SME-Fit Question

A range of barriers faced by SMEs that adopt AI via vertical SaaS are well-known and already extensively documented in qualitative SME pilot studies. A set of common organizational challenges is highlighted, including lack of in-house skills, data quality and governance problems, insufficient understanding of AI capabilities and of potential use cases within the firm, security and privacy concerns, and lack of ROI metrics, which were consistently cited [42]. Taken together, these constitute what can be termed the organisational readiness gap. Similar conclusions emerge from an analysis of the strategic barriers facing organizations in the public sector in terms of the adoption of AI technology, which finds that skills, data quality and governance, and scaling from pilot projects to production environments are all the most difficult challenges to address through just technology [61]. This cost sensitivity is explicitly included in the finding that SMEs tend to be more cost-sensitive than other enterprises and therefore cost of implementing and maintaining AI solutions in terms of the time and resource required is likely a major factor inhibiting adoption [59].

Some of these last hurdles are eased by the European regulatory framework. The EU AI Act was adopted by the European Union and lays down common rules for the placing on the market, and the putting into service and use, of AI systems and classifies them into four categories of risk (unacceptable, high, limited and minimal) [61]. To facilitate access to and uptake by SMEs, the European Commission has allocated a pilot fund of more than €100 million and promised additional measures within its 2024-2025 communications on European AI capacity [59, 61]. Whatever these measures, all studies concur that, regardless of the size of the enterprise, investment into data quality and governance is a prerequisite for all AI implementations [42, 58].

In summary, adopting the SaaS route removes the integration problem but not the problems of data preparation, governance or organisational change. It turns an engineering problem into a procurement problem; it remains the case that procurement requires some judgment as to whether it is a good workflow fit and how the vendor is going to handle its own firm's data.

4.6 Barrier-by-Barrier Synthesis: From Exclusion to Accessibility

The preceding sections have described the architecture of exclusion and the 2020-2024 transitions that dismantled each of its pillars. The evidence is now assembled into a barrier-by-barrier reduction matrix, the residual barriers are measured where the literature supports that measurement, and the design principles that follow from this synthesis are described. The synthesis is not celebratory: the same studies that document the reduction of barriers also document the persistence of others, and the chapter closes by drawing this distinction carefully rather than eliding it.

4.6.1 The Reduction Matrix: Where Barriers Have Fallen

As documented in Section 4.2, inference cost fell 1,000-fold between 2021 and 2024 [44], removing the capital barrier identified as the top obstacle [41]. On the demand side, a survey of 5,232 SMEs across seven countries finds that 30.7% are already using generative AI, with 65.1% of adopters reporting improved employee performance and 45.2% reporting cost savings [65]. An analysis of 11,429 European SMEs finds that AI-adopting firms show a 15.1 percentage-point higher probability of achieving revenue growth of at least 30% [66].

The second pillar of vendor dependence has been partially eroded by the emerging open-weight model ecosystem and by the edge-deployment evidence presented in Section 4.3. The integration barrier has been addressed in part by the no-code platform discussion in Section 4.4. The final pillar of the skills required has shifted to prompt engineering, data governance, and procurement judgement over ML engineering [65]. Section 4.5 shows the vertical SaaS solution which will eliminate this cost for many use-cases. Figure 4.5 and Table 4.1 provide summary and details of the four pillars of the pre-2020 baseline and their reduction, unevenly but measurably, by 2024.

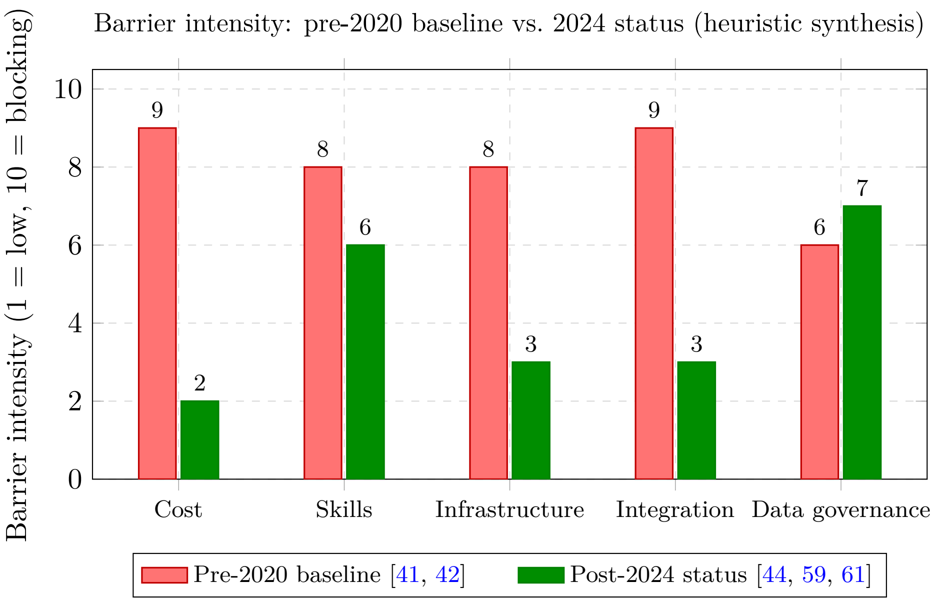


Figure 4.5. Heuristic synthesis of barrier intensity across the four pillars plus data governance, before and after the 2020–2024 ecosystem shifts. Cost, infrastructure, and integration have fallen sharply thanks to LLM pricing collapse, edge inference, and no-code platforms; skills and data governance have shifted in shape but remain elevated. Intensities are author-assigned on a 1–10 ordinal scale based on the qualitative evidence in the cited sources, not a direct quantitative measurement. Sources: Sysoiev [44], Horváth and Szabó [41], Závodná et al. [42], Florina et al. [59], OECD [61].

Barrier	Pre-2020 baseline	2024 enabler	Status
Cost	High capex; per-deployment training and tuning [41]	LLM-as-a-service at \$0.06–\$2/M tokens; opex model [44, 45]	Largely eliminated
Skills	Scarce ML engineers; SMEs out-priced in talent market [41, 42]	Shift toward prompt engineering, governance, procurement [59, 61]	Reduced; shifted in shape
Infrastructure	GPU clusters; in-house data centres; data sovereignty risk [41, 42]	API access; edge inference (EdgeShard 50% lower latency) [17, 47]	Largely eliminated
Integration	Bespoke engineering; months-long pipelines [41]	No-code/low-code; orchestration frameworks; vertical SaaS [49, 51, 54, 56]	Largely eliminated
Data governance	GDPR overhead; ad-hoc compliance [41]	EU AI Act framework; capAI; SME-specific maturity models [58]	Residual; reshaped

Table 4.1. Barrier-by-barrier reduction matrix. Cost, infrastructure, and integration have been largely eliminated by the 2020–2024 ecosystem shifts; skills and data governance have been reduced but reshaped, requiring organisational rather than technological mitigations.

A retail study calculates the value associated with this cumulative decrease. Generative AI is reported to be able to add \$240 billion to \$390 billion in value to retailers, representing a margin improvement of 1.2 to 1.9 percentage points [62]. Beyond the headline absolute value, the key indicator is that when companies adopt Gen AI, the barrier decreases and creates quantifiable value at the company level. Over 50 retail executives were interviewed who have either already begun piloting or scaling LLMs and Gen AI across their business, though only two of the respondents had Gen AI rolled out to their entire company, proving the challenge from pilot to production remains. In practice, the Amazon case serves as an example: according to a relayed Amazon report, in late 2023 Amazon launched AI-powered image generation tools for advertisers which boosted click-through rates by as much as 40%. In addition, AI-powered coding assistants are reported to be able to reduce the time spent on software engineering tasks by up to 60% [62]. The value potential of generative AI in retail is estimated to result in 5% additional sales and a 0.2 to 0.4 percentage point margin improvement in EBIT. This is also reflected in the cost of running an underlying API as well with GPT-4o, released in May 2024, which runs for half the price it took to run GPT-4 Turbo just one year ago, while industry observers are reported to anticipate further reductions of up to 80% within two to three years. Lindex is an additional example from Europe that provides some insight: in June 2023, the Swedish retailer unveiled the Lindex Copilot to aid retail staff in its physical locations by training it using information collected from Lindex sales and store data.

4.6.2 Residual Constraints That Survive the Reduction

Skills and knowledge continue to be the most commonly mentioned residual obstacle, documented across sectors and geographies. Estimates suggest that over 80% of AI projects do not succeed in reaching their objectives, double the rate of non-AI projects [61]. Furthermore, in a 2024 survey covering five countries, participants said the lack of in-house skills was the single biggest issue holding back AI adoption, with 60 percent of public-sector participants saying so. Likewise, a study carried out at the national level found that 70 percent of government agencies in the UK said skills were a barrier to their AI work. And in an online survey conducted in August-October 2024, which questioned 2,000 employees of the Australian Public Service about their experience using AI, 92 percent had never received any training in how to use the tech and only 16 percent felt comfortable using AI. Similarly, the report also notes that of the EU public-sector AI cases analyzed by the Public Sector Tech Watch observatory, 58 percent are still in the plan, pilot, or in-development phase. This means most of the cases are either still in an experimental phase or have not been implemented yet [61].

An ongoing difficulty is data governance. Only 59 percent of OECD countries have a public sector data strategy and there are very few countries with guidelines on how to set this [61]. The data on risk perception reveals a similar trend. A 2024 poll reports that 63 percent of public sector respondents believe that generative AI will undermine trust in institutions, both globally and within their country [61]. There are also different costs for skills sourcing, internally and externally. Bringing in contractors could cost three or four times as much as taking on public sector workers. This creates a problem for government whereby they are uncertain about up-skilling or re-skilling needs and the skills they should hire externally. There is also insufficient transparency to allow the government to understand what services that are being sold by vendors so that they can buy the right things or services at the right price. And finally policy maturity. Only two percent of the national AI policy that are recorded within the OECD AI database for National Policies and Strategies are risk-weighted as a way of moving away from a "one-size-fits-all" towards more data-driven and risk-based approaches.

These insights extend to developing economies as well. The *Digital Progress and Trends Report 2025* frames AI readiness through four foundational dimensions: connectivity (energy and digital infrastructure), compute (AI chips, data centres, cloud computing), context (data), and competency (skills); strengthening these foundations is identified as a precondition for inclusive AI adoption [60]. The same report observes that, alongside the visible foundation-model race, many developing economies are deploying compact, lower-cost AI solutions tailored to local needs and budget constraints. Three complementary policy directions are also recommended: strengthening data-governance frameworks, advancing regulatory and institutional reform, and promoting digital-era skill formation.

The residual claims are reinforced by data on European SMEs. Forty-one percent of SMEs consider the GDPR cumbersome, while the fear of penalties or proceedings discourage AI implementation in areas using data [59]. Over 50 percent of SMEs employees have insufficient digital competencies to perform AI tasks. Several recommendations for SMEs are put forward. A gradual approach to implementing AI is advised, starting in specific business areas rather than large-scale, all-at-once deployments. Also recommended are effective data governance that preserves data protection by design while building trust and security, training to equip employees with new digital skills and technologies, and partnerships that build AI collaboratively through cooperation between SMEs, AI providers, research institutes, and government-funded innovation centres. Some policy recommendations reinforce this picture: grants, subsidies and tax schemes for AI investments, regulatory sandboxes to experiment with AI in controlled environments, upskilling with universities and private technology firms and international co-operation to build standards of AI regulation, cross-border and technology transfer [58, 59].

4.6.3 Design Principles for the Operational Playbook

The synthesis yields five design principles. They bridge the empirical analysis presented above and the strategic frameworks developed in the chapters that follow.

To begin with, the high cost barrier of AI has, for the most part, been removed for the majority price-sensitive SME applications: the LLM-as-a-service unit economics enable affordable experimentation for firms of any size [44, 62]; thus, AI investments no longer need to be seen in the light of capital expenditures. Furthermore, the integration cost is low enough that domain knowledge, not engineering skills, is the primary constraint in no-code and embedded productivity settings [49, 52]. Third, vendor lock-in remains a material concern for data-sensitive processes; it is mitigated through the use of open-weight models alongside infrastructure design that supports multi-vendor switching [17, 47, 48]. Fourth, the residual barriers are organisational rather than technological, requiring upskilling, data governance, and change management to bridge the gap from experimentation to production [42, 58, 59, 61]. And fifth, while there is a high level of variance in the SMEs' readiness to adopt AI, there is no universally applicable approach to the adoption problem [67]; therefore, the next section provides a scaffolding to assist decision-making rather than a prescriptive set of rules for the SME to follow, beginning with a maturity model in the subsequent chapter.

The main references for this chapter are listed in Table 4.2 by types of evidence. This categorisation makes obvious an imbalance in sources: the price-collapse and infrastructure findings in Sections 4.2 and 4.3 come from refereed research, whereas many of the high-profile market and value numbers in Sections 4.5 and 4.6 come from consultancy and vendor reports and should be interpreted with caution.

Table 4.2. Summary of the principal sources informing this chapter, by evidence type.

Source (author/year)	**Type**	**What it supports**
Sysoiev (2025) [44]	Peer-reviewed	LLM inference price collapse \$60→\$0.06; price tiers; startup formation and VC flows (§4.2)
Feuerriegel et al. (2023) [45]	Peer-reviewed	Three-layer model/system/application framework for generative AI (§4.2.1)
Lins et al. (2021) [18]	Peer-reviewed	AI-as-a-Service; capex-to-opex framing (§4.2)
Luitse (2024) [17]	Peer-reviewed	Cloud concentration (68% AWS/Azure/GCP); three-layer stack; lock-in (§4.3.1)
Pereira et al. (2024) [48]	Peer-reviewed	Open-weight model value proposition: transparency, cost, customisation (§4.3.2)
Zhang et al. (2024) [47]	Peer-reviewed	EdgeShard collaborative edge inference; 50% lower latency, 2× throughput (§4.3.3)
Berberi et al. (2025) [51]	Peer-reviewed	MLOps and orchestration landscape (Airflow, Kubeflow, Prefect) (§4.4.1)
Hamdhana (2024) [49]	Peer-reviewed	No-code/low-code example (Flowise with RAG) (§4.4)
Engstrom & Gelbach (2021) [53]	Peer-reviewed	Legal-tech vertical example (§4.5.2)
Thota (2025) [56]	Peer-reviewed	CRM/Salesforce Einstein multi-tenant AI; accuracy gains (§4.5.2)
Bäumer et al. (2022) [57]	Peer-reviewed	Manufacturing machine-translation SME case from ∼5,000 PDFs (§4.5.2)
Horváth & Szabó (2019) [41]	Peer-reviewed	Pre-2020 baseline: cost, skills, infrastructure barriers (§4.1)
Závodná et al. (2024) [42]	Peer-reviewed	22-expert CZ+AT pilot; four canonical barriers; SME population (§4.1)
Florina et al. (2025) [59]	Peer-reviewed	European-SME residuals: 41% GDPR, competency gap; recommendations (§4.6)
Eberhardsteiner (2024) [58]	Master's thesis	SME data-governance and capAI maturity (§4.6, table)
OECD (2025) [61]	Institutional	Residual skills and data-governance; AI-system definition; EU AI Act risk tiers (§4.6)
World Bank (2025) [60]	Institutional	Four-C readiness framework (connectivity/compute/context/-competency) (§4.6.2)
Gartner (2024) [46]	Consulting/vendor	30% of API-demand growth AI-driven by 2026 (§4.2.3)
Ernst & Young (2024) [50]	Consulting/vendor	Low-code/no-code governance ver-

Chapter 5

Categorization of SME Problems from AI Use Cases

5.1 From Heterogeneous Problems to AI-Ready Categories

The list of structural and functional challenges confronting small and medium-sized enterprises assembled in Chapter 3 is illuminating, but on its own it does not tell the reader which problems are readily solvable today, which require major change, and which families of AI tools are likely to apply in each case. It is therefore the task of this chapter to organise that problem space into a taxonomy of categories on which the remainder of the thesis can rest, and to define the dimensions, criteria, and empirical evidence that justify those categories.

This section argues that categorization ought not be treated as an aesthetic consideration, but as a prerequisite. Without it, one cannot hope to create a Chapter 6 that pairs solution patterns to problems with a modicum of accuracy, nor a Chapter 7 that produces an actionable implementation guide rather than a vague set of platitudes. Without a taxonomy, all the problems are essentially alike to the manager who comes to AI cold. With a taxonomy, problems become commensurable, comparable and can be grouped with families of technical solutions.

5.1.1 Why a taxonomy is needed before solutions

Between 2022 and 2025, the literature on AI adoption in SMEs has grown substantially, yet none of the three main evidence streams, academic, institutional, or consulting, offers a usable problem-level taxonomy. A systematic review of 64 peer-reviewed papers identified 27 distinct adoption challenges but, by design, does not classify the problems themselves [29]. The 2025 G7 discussion paper supporting the proposed SME AI Adoption Blueprint [36] classifies adopters by digital maturity and usage complexity, useful for policy targeting, but it does not categorize the problems an individual company faces; the documented adoption gap between large firms and SMEs is addressed function by function, not problem by problem [68]. Consulting sources such as the NASSCOM-EY AI Adoption Index [69] and the State of AI 2026 [70] document uneven progress and highlight that the framing of AI shapes its application, but neither is organized around a taxonomy of SME problems. It is this classification, simultaneously considering business function and AI capability, that the present chapter provides.

5.1.2 Methodological approach: triangulation across three sources

First, peer-reviewed research. The primary reliance is on the systematic review [29], with additional support from papers on specific kinds of use cases, e.g., conversational AI in SMEs, and the subsection on AI process automation. Academic research reveals what has been studied empirically and what has not, and identifies where the taxonomy ought to be constrained: to what has been shown, rather than where the vendors are most excited.

The second type of source is that of institutional reporting. The Discussion Paper on SME adoption of AI among G7 countries [36], the Taxonomy for the European AI Ecosystem [71], the Annual Report on European SMEs, and the D4SME surveys offer quantitative coverage at the national and sectoral levels. Institutional sources are important because they provide methodological transparency by reporting their sampling and definitions and are not likely to overstate adoption with the aim of bolstering a commercial product or service. For this latter purpose, the same taxonomy is particularly useful, as it specifically attempts to align AI use cases to organizational and sectoral dimensions through an ecosystem approach to the structural framework analysis.

The third source is consulting research from the major professional-services firms [72]. Consulting reports vary widely in scientific accuracy, but they possess two things that academic and institutional research typically do not: insights gained through firsthand observations of actual consulting projects, and industry classifications that map to categories in which consulting firms have been hired and have actually done work. This section most frequently relies on the consulting research reported in the State of AI 2026 report [70] and the NASSCOM EY index [69], supplemented by other consulting research when it provides more clarity for certain categories.

These three evidence sources can be used together to reinforce the validity of the taxonomy, in response to common critiques that might be raised: It is not simply an academic abstraction because it is rooted in reality by the practice-oriented consulting sources; it is not simply a vendor-centric feature list because the academic and government sources filter the features that are included; it is not solely a summary of high-level concerns because the practice sources are the most closely tied to actual deployment and usage.

5.1.3 Inclusion criteria for a problem to enter the taxonomy

Not all SME problems qualify for inclusion. Four criteria, each grounded in the evidence cited above, determine entry.

The first criterion for each category is that the issue is a recurring problem in multiple SMEs, which can be observed as part of the structural issues described in section 3 of this paper or within the use case statistics gathered from Eurostat and the SBA and not as an isolated occurrence in a specific niche. An analysis of the literature reveals that the majority of mentioned barriers and use cases focus on a small range of activities that represent the backbone of this categorization [29].

The second criterion is empirical evidence that AI is being adopted in practice in the present day; this will be at modest levels, but SMEs will have demonstrated a positive uptake already. The figures provided in Chapter 3 in Eurostat data and the evidence presented in Section 5.2 below will set the bar. The categories which may show signs of enthusiasm but not yet demonstrated uptake fall below this threshold and are excluded from the primary taxonomy, as discussed in Section 5.2.5.

The third criterion is gap-fit with at least one barrier identified in Chapter 3 or Chapter 4. A category is included in the taxonomy where a known AI capability solves a documented SME pain point. If a category defines a new SME pain point that was not identified in the preceding chapters, then that category is excluded. This ensures the taxonomy is closely related to the upstream chapters and does not look like a completely new list of vendor desires.

A fourth condition is coherence with management taxonomies. Categories must fit the mental model used by SME managers: functions, customers, decisions, knowledge, documents. In cases where a technically elegant AI capability breakdown conflicts with how an SME owner structures tasks, priority is given to the management perspective. This is why the chapter will be usable in Chapter 7.

5.1.4 Three dimensions of categorization

While this chapter is structured by problem areas instead of technologies, these categories are not one-dimensional. Three underlying dimensions underpin each of them and appear in each subsequent section.

The first dimension is problem area. This is the functional anchor and what gives the taxonomy its surface labels: document and process automation, customer service and engagement, marketing and sales, knowledge management and internal search, decision support and analytics. These five labels are the categories proper. This is the way SME owners would naturally identify the place where their work takes place.

The second dimension is the prevalence of AI capability, based on the four-way split in Chapter 2: perceptual, predictive, generative, agentive. Each dimension has one or two prevalent capabilities, though not limited to that. Document automation is driven by perceptual and generative capabilities, customer service by generative and agentive capabilities, knowledge management by retrieval augmentation and embedding search, decision support by predictive and analytical capabilities. It matters for the name of the prevalent capability because this limits what a real life technology stack and skills set can look like.

The third axis is adoption depth, as spelled out by the SBA survey mentioned in Chapter 3. Certain buckets allow for “plug-and-play” onboarding, in which an SME can adopt a SaaS product with relatively little integration. Others necessitate much more extensive changes in business process, data infrastructure or governance. The taxonomic bucket for Customer Service (Section 5.4) is the most plug-and-play of all, consistent with the finding that generative AI adoption in SMEs is highest in customer-facing functions [65]. By contrast, Decision Support (Section 5.7) is among the deepest, in most

SME settings. The dimension of depth matters for purposes of sequencing, an ordering summarised in Section 5.8 and developed into a staged plan in the playbook of Chapter 7.

The following subsection presents the convergence of evidence across problem areas. Table 5.1 at the end of Section 5.2.5 summarises the resulting five-category taxonomy. Generative AI spans categories 1, 3, 4, and portions of 2, which helps explain why the advent of accessible language models from 2022 has lowered the barrier to entry for a great number of categories simultaneously rather than just for one or two categories individually [36, 65].

5.1.5 Five categories and the road through this chapter

Sections 5.3 through 5.7 will cover the five categories of document and process automation, customer service and engagement, marketing and sales, knowledge management and internal search, and decision support and analytics. The paragraphs on these categories will all follow the same structure: begin with a definition and outline of the category, then ground it empirically in the data from Chapter 3, and evidence from OECD and consulting reports, introduced in Section 5.2, then describe the types of problems which are part of the category, then the specific barriers which SMEs face in adopting solutions for the category, and end with a link forward to the types of solutions covered in Chapter 6.

The chapter then concludes by synthesising the five categories in a closing summary (Section 5.8), reading them as an ordered sequence of ascending barrier height, identifying the cross-cutting constraint that binds across all of them, and making the transition to Chapter 6, where these same categories become the entry points into solution patterns and tool families, and to the implementation playbook of Chapter 7.

The taxonomy presented in this chapter is intentionally not exhaustive or fixed. It is a principled but scope-bounded structuring of the problem space, grounded in the four inclusion criteria of Section 5.1.3, that serves the operational guidelines of the present thesis. As agentive AI develops and its integration patterns mature beyond 2026, the categorical delineations can be expected to shift accordingly. What should persist is the practice of first defining before prescribing, and of first defining by multiple evidentiary sources before assigning an appropriate action.

5.2 Empirical Foundation: SME Digital Maturity and Process Priorities

Taxonomy is only ever as good as the factual basis on which it is built. Before turning to these categories, this section sets out in a little more detail the picture of the level of digitalisation and AI adoption amongst SMEs that justifies the categories chosen for this study, and explains why others have been rejected. There are three main sources: the Annual Report on European SMEs of 2023/2024 [68], the D4SME Survey of SME Digitalisation of 2025 [73], and the discussion paper of 2025 on SME AI adoption for the G7 [36]. Combining these with the earlier and more comprehensive 2021 study of the Digital Transformation of SMEs [28], research on the drivers of SME digital adoption in

Europe based on 15,346 SMEs [74], the NASSCOM-EY AI Adoption Index for India [69], and survey-based findings regarding what SMEs are actually asking for from their digital service providers [75].

5.2.1 The shape of the SME population in Europe

European policy discussions tend to regard SMEs as a sort of leftovers to the real economy, which is, in demographic terms, European. The 2023/2024 SME Performance Review shows small and medium-sized enterprises form the majority of companies in the European Union's non-financial business economy [68]. It is also year-over-year population of SMEs is reviewed, but the consequence of this review is that a rather varied and not-so-tech-savvy population is subject to technological progress, which up until 2022 was developed mostly by and for big business enterprises. The G7 evidence reports that the adoption gap between large firms and SMEs is apparent across all G7 countries [36]; the EU data echoes the same message, just differently.

Shifting from workforce size to digital engagement, the 2025 D4SME Survey of 10 OECD Member countries, namely Australia, Canada, France, Germany, Italy, Japan, Korea, Spain, the United Kingdom and the United States, demonstrates that SMEs gain from digitalisation through multiple channels but encounter consistent challenges in using new technology and restructuring their own business practices [28, 73]. The D4SME survey was conducted with four thematic priorities in mind, three of which directly overlap with the classification in the present chapter: adoption of AI and generative AI technologies, digital practices for resilience and sustainability and digital security practices. The conceptual framework of the D4SME is important because it shows that adoption of AI does not constitute a singular, one-time decision made by an SME, but rather consists of layered decisions regarding which parts of their work to automate, augment and leave to people.

The longer-form 2021 report The Digital Transformation of SMEs builds on this [28]. It confirms that digitalisation enables SMEs to enjoy scale economies and lower transaction costs, a set of competitive advantages until now reserved only to large firms. While adoption has already occurred to some extent across the entire portfolio of digitally-able tasks, it is most advanced where the threshold for entry is lowest: customer-facing channels and basic productivity tools. Other areas that require closer integration of the digital layer with operations or production have barely seen uptake. While this asymmetry is a single indicator, as a collective, it helps inform why the taxonomy in the current chapter makes distinctions across categories where adoption is already meaningful versus early-stage.

5.2.2 What drives adoption, and what fails to

For a majority of SMEs, digitalisation is a prerequisite for adopting AI. The question thus is: what determines digitalisation? An ordered logit regression model was tested on 15,346 EU and non-EU SMEs with a technology-organization-environment (TOE) framework [74, 76], identifying three main factors. These findings contribute to the taxonomy in the following ways. First, the most critical factor, the so-called technology context, is the

IT infrastructure and the existing set of digital tools. Firms already relying on cloud, accounting systems, or e-commerce tools are more likely to adopt digital technologies further along the digital continuum. Second, the existing levels of innovation in the firm, which captures the firm's organizational propensity for novelty, is also critical. Third, the third pillar of the TOE, which is the environmental context, has little impact on the digital technology adoption decision. The analysis claims that this implies a greater sensitivity of the SME decision to internal factors rather than external ones like external grant programs or industry-level innovation activities.

So, for this chapter, this means that for any category where SMEs have already built digital muscle, such as with their customer-facing technologies, AI adoption can be expected to happen faster compared to categories where the underlying digital base is not as robust. As an example, decision support and analytics depend on the existence of historical structured data to be functional and useful, and so SMEs that lack existing ERP or CRM systems would have a much longer journey to catch up with similar-sized firms that already have these systems.

5.2.3 Where AI adoption sits today: G7 evidence

The December 2025 G7 discussion paper [36] provides the latest dedicated evidence on SME AI adoption rates across G7 economies. Its central finding is that large and continuing disparities persist between SMEs and large companies in all G7 member countries, while a small group of SMEs has progressed beyond experimentation to productivity-relevant AI use. The paper classifies adopters along three dimensions, digital maturity, AI usage complexity, and scale, and will form part of the G7 SME AI Adoption Blueprint.

That taxonomy of adopters is a closely related conceptual framework; however, this system classifies firms by their use of AI rather than what they are using AI to solve. In that sense, the two are complementary. A firm categorised as an advanced user may be in just one of the five problem categories discussed in the present chapter. A firm categorised as a starting user might be involved in two of them already (albeit at lower levels of maturity). That classification will be most useful for cross-country comparisons, while the taxonomy in the present chapter will be more useful for individual SME decision-making.

5.2.4 What SMEs actually ask for: priorities from survey evidence

International organizations' statistics depict the supply side, including policy. They show us where SMEs invest and what is missing. But they do not reveal the priorities that SMEs have regarding the investments that they are prepared to make on technologies. This gap has been filled through interviews with SME owners and executives in three European countries that examine what matters the most to them with their banking and non-banking service providers [75]. Beyond banking, the findings offer an important perspective on preferences: SME executives are largely split into three categories: around 20 percent prefer digital-only channels, 20 percent prefer purely in-person service, and the majority of firms prefer some mix of both types of access. What they want is advice and support with environmental, social, and governance reporting, with tax, with finance and with the digital technologies they increasingly need to operate efficiently and effectively. They would like access to such services and tools through digital channels in particular.

This confirms the implicit assumption underpinning the Customer Service category in Section 5.4 and the document automation category in Section 5.3: that SMEs do not eschew technology and are not pursuing it for the sake of technology alone, but are rather requesting solutions to particular pain-points, and that these solutions are preferably bundled with solutions they already have access to. The finding that 53 percent of SMEs are satisfied with their bank's digital platforms, whereas 34 percent take some of their business to new digital entrants [75] is instructive of how they treat AI, with SMEs readily accepting AI when bundled within the service they already use, while they will take their business elsewhere when the incumbent solution does not offer it for an acceptable price or does not offer the same capabilities. For the taxonomy, this finding validates why categories 1 and 2 (document automation and customer service) exhibit the lowest skills barriers: SMEs prefer solutions bundled into tools they already use, which is precisely the SaaS-first deployment model that characterises these two categories.

5.2.5 Convergence on five problem areas

Overall, the empirical sources lead to organizing the problem space into five groups rather than four or six. Statistics on the SME population [68] and findings showing which capabilities get digitalized first [28, 73] both corroborate that customer-facing use cases and back-office process automation are where SMEs have made the most progress. This underpins categories one and two as the first two taxonomy groups most pertinent today. The evidence shows that existing digital maturity is the greatest factor determining adoption. Since many SMEs already have a variety of third party platforms to market and sell to customers, this validates a third group with high SME activity for marketing & sales. The G7 report [36] and the accompanying survey [75] show that knowledge work and decision support are where the SME vs. large firm divide is most pronounced, justifying these as groups four and five for the purposes of this chapter, since these are the groups where closing the gaps are most important. Data from the India-based index [69] confirms the five-group partitioning is not just a European policy artefact, but something that can be observed with other large SME markets.

The resulting five categories, listed in Table 5.1, are: (1) Document and Process Automation, covering invoice extraction, contract review, and administrative workflow; (2) Customer Service and Engagement, encompassing chatbots, ticket routing, and voice agents; (3) Marketing and Sales Optimization, spanning content generation, lead scoring, and personalisation; (4) Knowledge Management and Internal Search, addressing semantic search, question-answering over company documents, and expertise location; and (5) Decision Support and Analytics, including demand forecasting, churn prediction, and anomaly detection.

The choice of these five categories, and the exclusion of alternatives, rests on three arguments. First, every category satisfies all four inclusion criteria defined in Section 5.1.3: each is a recurring problem across multiple SMEs, each has demonstrated empirical AI adoption, each maps to at least one barrier from Chapter 3 or Chapter 4, and each aligns with how SME managers structure their work. Second, candidate groupings that were considered but rejected include supply chain optimisation and human resource management. Supply chain AI adoption remains concentrated in large firms with complex logistics networks [36]; the Eurostat data show negligible SME uptake (logistics at 6.1% of AI-using firms) [77], falling below the empirical adoption threshold. Human resource management, while a real SME pain point, lacks a distinct AI capability profile that would differentiate it from document automation (for CV screening) and knowledge management (for onboarding); including it would fragment the taxonomy without adding analytical power. Third, the five-group structure is parsimonious: it is the smallest number of categories that covers the documented clustering of SME AI use cases across academic, institutional, and consulting sources, without collapsing meaningfully different barrier profiles into a single bucket.

The empirical foundation does not provide a prioritization on what category an SME ought to start with. That is a matter of company-specific circumstances, summarised in Section 5.8 and worked out in detail in the playbook of Chapter 7. The empirical foundation simply indicates that the five categories are strongly supported by evidence as being the places where AI use cases are currently clustered, none of which is the creation of a vendor, and that the difference between SMEs and enterprises varies in a systematic way across the five.

Table 5.1. Summary of the five SME AI-ready problem categories, with primary AI capabilities and representative use cases. Derived from the triangulation of academic, institutional, and consulting sources discussed in Section 5.2.

Category	**Primary AI capability**	**Representative use cases**	**Key SME barrier**
Document & Process Automation	Perceptual + Generative	Invoice extraction, contract review, form processing	Data quality, legacy integration
Customer Service & Engagement	Generative + Agentive	FAQ chatbot, ticket routing, voice agent	Training data scarcity, brand voice
Marketing & Sales	Generative + Predictive	Content creation, lead scoring, personalization	First-party data, attribution
Knowledge Management	Generative (RAG)	Semantic search, Q&A on docs, expertise location	Document silos, privacy (GDPR)
Decision Support & Analytics	Predictive	Demand forecasting, churn prediction, anomaly detection	Data history, analytical skills

5.3 Category 1: Document and Process Automation

The first category in the taxonomy is document and process automation. This is the set of problems that are repetitive document management tasks, including: manual data extraction from unstructured/semi-structured sources, invoice processing, contract management, and general admin tasks that consume a significant amount of small business operating time. This is where SMEs have the strongest urge to solve, where the technology stack is ready, and where the time from POC to full scale is usually minimal [29, 70].

5.3.1 Definition and scope

For the purposes of this thesis, document and process automation encompasses the end-to-end handling of business documents that arrive as unstructured or semi-structured inputs and must be converted into structured, actionable data within existing enterprise systems. The category spans extraction, generation, routing, and exception management, each of which is detailed as a distinct sub-type in Section 5.3.2.

This classification is unique in that it brings together two separate, longstanding automation movements. One is Robotic Process Automation (RPA), the technology used for the repetitive, rule-based parts of the workflow, parts that take in structured data and act upon defined screens, the second is the new wave of cognitive automation technologies, built on technologies such as optical character recognition (OCR), document understanding, and language models, which deal with those parts of the workflow that are not defined by rules. These two movements are coming together, what is sometimes termed Intelligent Document Processing, and this is what is making the category relevant today to small and mid-size enterprises that could not previously afford either a pure RPA solution because the volumes were too low, or pure cognitive automation because the technology was too expensive. The category is now at a price-performance point that fits the SME budget [18, 44].

The reality-based evidence comes from Chapter 3, which outlines the scale of the administrative burden. The German survey of small and medium enterprises (SMEs) found the average SME dedicates seven percent of working hours to administrative activities, the equivalent of approximately 32 hours per month per business, and ranked bureaucracy as the single largest perceived risk to domestic competitiveness [33]. That number becomes the pivot point for value: each hour spent on document and process-related automation is an hour that returns to the hands of a small business owner or their team, freeing them to work on revenue-driving activities. This category does not sell the technology, it sells bandwidth recovery.

A systematic review affirms the empirical significance of this category in scholarly literature [29]. The review notes that process automation and automation of administrative tasks are business processes where AI applications for SMEs are most prevalent and the disparity between large and small companies is smallest. A PRISMA-protocol analysis of 64 articles determined that the barriers most frequently encountered in the adoption of AI for SMEs are knowledge gap, financial costs, and inadequate infrastructure; and that process automation is one of the limited domains in which the second and third of the above-mentioned constraints have been alleviated due to the software as a service cloud-based business model. The knowledge barrier remains, albeit in this category the hurdle is one of awareness instead of feasibility.

A 2026 report on the state of AI in the enterprise identifies process automation as a predominant use case of AI in the organizations surveyed, focusing mainly on efficiency-driven use cases [70, 78]. It also draws a clear distinction between AI experimentation and genuine transformation, noting that while many organizations concentrate on efficiency-driven uses of AI, a smaller group takes the further step of reimagining business models, offerings, roles, and ways of working. For SMEs, the efficiency use cases of document and process automation are not shortcomings, they are the right places to start. The reimagining work happens downstream, on a path summarised in Section 5.8.

5.3.2 Sub-types of problems inside the category

Within document and process automation, four sub-types of problems recur across the SME population.

The first one is invoice processing automation. SMEs that process anything from fifty to a few thousand invoices a month typically cannot justify the cost of dedicated accounts payable software featuring OCR, and the alternative is manual data entry, which is neither very reliable nor quick. But now AI-based invoice processing is readily available in the form of SaaS products, offering OCR and language-model-based field extraction, supplemented by rule-based validation. So now an SME can extract data including the supplier name, invoice number, date, line-items, amounts, and tax codes from a PDF or photo, and inject it into their own existing accounting system without needing a dev in the loop. A 2022 SME order-entry process case study shows that same pattern at a more granular level, demonstrating that RPA combined with AI could be productive even when the process "involves a small number of repetitions, but a high level of individual arrangement" [79].

The second sub-type is contract analysis and review. SMEs will often sign contracts that they do not have the staff to review carefully and that use template agreements, the use and drift of which they do not readily track. AI-powered contract analysis tools, built originally for big law departments, are becoming more available in smaller form factor for SMEs and can flag deviations from a reference template, extract obligations and dates, and draw a user's attention to specific clauses [53]. The systematic review notes that legal and contract-related AI applications remain among the less studied areas for SMEs [29], indicating a sub-type that is less mature than invoice processing but growing.

The third sub-type includes form-processing and data entry from semi-structured sources. The data in forms received by small businesses, e.g., customer information, employment records, supplier specs, compliance filings, has varying enough structure that templating does not work; yet an identical intelligent document processing platform used to process invoices can also be reconfigured for such forms; and the cost of that reconfiguration has fallen with the emergence of subscription-based, no-code IDP platforms accessible to non-specialist users [18].

A fourth sub-type is workflow automation augmented by AI. The typical RPA pattern (in which a software robot automates the use of a series of applications, to transfer information and activate events) becomes highly customizable when the robot can invoke a language model for handling exceptions, rephrasing, and other edge conditions otherwise requiring a human. The case study shows this clearly: the RPA+AI application examined for an order-entry process provided quantifiable productivity gains in a workflow that, according to the SMEs, was too customised for standard RPA [29, 79].

Table 5.2. Sub-types of document and process automation problems, with their primary AI capability, data requirements.

Sub-type	**Primary AI capability**	**Data requirement**
Invoice processing	OCR + LLM field extraction	Low: existing invoices suffice as training input
Contract analysis & review	NLP + template matching	Medium: reference templates and clause libraries needed
Form processing & data entry	Intelligent Document Processing (IDP)	Low: existing forms, reconfigurable IDP platform
Workflow automation (RPA + AI)	RPA + LLM for exceptions	Medium: process mapping and integration with existing systems

5.3.3 SME-specific barriers in this category

Although automation of documents and processes has the lowest residual barrier height - consistent with its status as the most prevalent SME AI use case with the smallest large-firm disparity [29, 77] - there are nonetheless specific remaining barriers worth noting.

The first is data quality. SMEs will receive documents in multiple languages, multiple formats, and often in the form of poor-quality scans or photographs. AI extraction tools will work beautifully on clean, single-language data but degrade significantly on heterogeneous inputs [29]. The only mitigation is to invest in document quality at the source (either by training staff and/or suppliers to follow a standard, or by selecting tools that have been trained on noisier data).

The second major hurdle is how well the tool integrates with a company's existing accounting and enterprise resource planning (ERP) systems. SMEs, in particular, often lack modern application programming interfaces (APIs) for these systems, which means the benefits of automated extraction could be lost if the extracted data still have to be manually inputted by a human. This has been addressed by the emergence of integration platforms and pre-built connectors offered by some of the leading SaaS providers, and the practical advice is to verify that connectors are available before using a tool.

The third hurdle involves the audit trail and accountability. Automated extraction introduces a type of error that humans do not introduce and that SMEs are not prepared to defend in an audit. Today's best-in-class tools provide for human-in-the-loop review of low-confidence extraction. The key in making a productivity tool into a defensible compliance process is the disciplined use of that review.

The systematic review [29] affirms that data quality, integration, and lack of in-house knowledge to evaluate AI are common themes in the SME literature and occur especially often in papers examining automation applications. These are thus genuine barriers, but they are addressed with current technology. This is why they constitute a distinct category from the ones discussed in the next two sections that will involve larger investments.

5.4 Category 2: Customer Service and Engagement

The second includes any interaction with the customer after they have become the firm's lead: calls, support tickets, appointments, FAQs, follow-up messages, and escalations. In SMEs, the category straddles an awkward line: the work often falls to a founder or a small team of two or three people; the volume of work rarely justifies an entire customer service tech stack; the costs of a missed call or a delayed response is felt immediately as lost revenue or a churned customer. The category has emerged as one of the most obvious applications for AI in SMEs exactly because there now is a price point that allows an otherwise unaffordable problem to be solved [65].

5.4.1 Definition and scope

This taxonomy includes four broad workflows when it comes to customer service and engagement: interactive question-answering, in which a customer has an active dialogue with a firm to, say, get information about a product, find out the status of an order, book an appointment, or fix a straightforward problem; ticket management, which entails sorting incoming requests, prioritizing them, and forwarding them to the right person or queue; sentiment-aware response, which involves reading the messages that customers send in order to assess their tone and urgency before processing them, ensuring that grievances are forwarded to the appropriate people immediately; and proactive engagement, when the firm initiates contact with customers, responding to cues that indicate what they might need based on their behavior as opposed to having customers reach out to the firm.

The primary AI features are generative and conversational, in that a generative model produces the natural-language response, a conversational agent controls the multi-turn dialogue, and a classification model directs or prioritizes messages. When the customer pool is large and there is abundant historical data, predictive models can be leveraged to identify customers who are at risk of churn, but, again, it is unlikely this predictive model capability will be a primary feature for most SMEs (but a feature that can be added at a later point in time).

Two empirical findings render this category particularly pertinent to SMEs. The first is the difference between large-firm adoption rates and small-firm adoption rates of chatbots and customer service automation, as documented by the results of the U.S. SBA survey discussed in Chapter 3, in which small firms trailed by about seven percentage points their large-firm peers in their adoption of chatbots [38]. That is a gap that matters because customer service is an activity that is characterized by particularly positive marginal cost conditions for automation [65]: A single chatbot can field scores of routine customer questions that would otherwise divert the owner from the tasks that only the owner can perform.

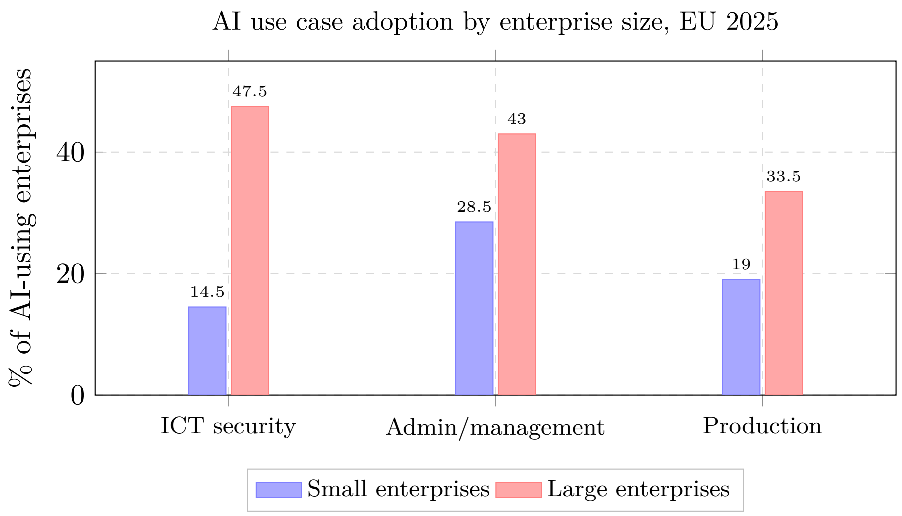


Figure 5.1. Share of AI-using enterprises deploying AI within each functional area, by firm size. Across all three use cases, large enterprises lead small enterprises by a substantial margin (+33 p.p. in ICT security, +14.5 p.p. in admin/management, +14.5 p.p. in production). Source: Eurostat [68].

The second trend relates to the evolving standards for customer service. The AI Playbook for India's MSMEs indicates that small businesses are currently at different phases of their digital transformation. And it identifies customer service automation as one of the viable options for those early on in the digitalization path [36, 80]. Further, it suggests that a small business taking on AI effectively needs to follow a structured approach, in which the stakeholders involved address the difficulties they encounter as they move from understanding AI to actually using it. Customer service is seen as one of the more accessible entry points, with the caveat that integration with existing customer data is the hidden cost that determines whether a deployment delivers on its promise.

In line with the larger market, the State of AI in the Enterprise 2026 survey shows that the majority of enterprise customer-facing use cases are already AI-powered, with the frontier now shifting to agentic AI systems capable of performing complex, multi-turn customer conversations autonomously [70, 72]. This framing for agentic AI is significant because it is indicative of where customer service is headed, moving away from AI that answers the question and towards an agentic AI that completes all the steps, including booking an appointment, updating an inventory system, and conducting follow-up post-service, within a single conversation. Even though SMEs will not be at the forefront of this development, they will be uniquely positioned to leverage agentic AI once low-cost SaaS products become available [36, 72].

Four sub-types of customer service problems recur across SMEs.

One specific sub-type is the use of chatbots and virtual assistants for FAQ purposes. Imagine, for example, an SME website that always gets the same few basic inquiries prior to purchase: store hours, returns policy, product compatibility, and delivery dates. By training a chatbot with that firm's FAQ and product information, the chatbot is able to address those questions automatically and hand it off to a human staff member when the customer has something that the bot cannot answer. This is exactly what was observed in a 2019 study of e-commerce adoption in the small and medium-sized enterprises of Malaysia, which developed a conceptual framework that used a tri-theoretical foundation based on the UTAUT and the TOE to explore SME intention to adopt intelligent conversational agents [81]. Its findings suggest that the decision is based on the degree of perceived ease of use combined with expected usefulness, in addition to the technology-organization-environment factors that determine whether or not the SME actually has the means of deploying the virtual agent.

The second sub-type involves ticket classification and routing. Customer service messages come in through a variety of mediums (email, online web forms, social networks, in some cases WhatsApp) and therefore need to be classified into the appropriate category and directed towards the person best qualified to handle them. Even a basic classifier that can distinguish between, for example, "order status", "complaint", and "general query" would help an SME, where perhaps just two or three people are looking after their customer service queries, by lightening their mental workload and speeding up their response times.

The third sub-category is sentiment analysis on customer feedback. SMEs gathering reviews, responses from surveys, or other unstructured feedback through their own channels can use sentiment classification to detect complaints. Although this use case is also closely related to marketing, the operational logic applies to the customer service function: a publicly-facing, strongly negative review could harm a company's reputation if it goes unanswered for long, and sentiment classification affords a lean team an early warning it would not otherwise have.

A fourth sub-type is voice enabled customer interaction, inbound call taking, appointment booking via voice. This sub-type is the newest to be available to SMEs [80], with the emergence of voice AI models that can integrate a speech recognition, language understanding, and speech synthesis into a low latency stack. Voice AI can now substitute human answering of phones during after-hours and peak times for some SMEs in service businesses where the phone remains the predominant channel (i.e. medical and legal practices, repair shops).

A 2023 exploration of conversational AI for the SME sector covers a range of these sub-types, combining both benefits and issues using a literature review and thematic analysis. That research looks at chatbots, virtual assistants, and automated response systems, concluding that conversational AI can significantly benefit SMEs while cautioning that it should be regarded as an augmentation of existing operations rather than a complete replacement [82]. This aligns with the current thesis: SME customer service functions are augmenting an under-resourced area of support rather than replacing a fully-staffed department.

Table 5.3. Sub-types of customer service and engagement problems, with their primary AI capability, data requirements.

Sub-type	**Primary AI capability**	**Data requirement**
FAQ chatbot & virtual assistant	Generative (LLM)	Low: product docs and FAQ seed examples suffice
Ticket classification & routing	Classification (NLP)	Low–Medium: historical tickets for training
Sentiment analysis on feedback	Classification (NLP)	Medium: review corpus and survey data
Voice AI (inbound calls, booking)	Speech-to-text + LLM + TTS	Medium: call scripts, business rules, telephony integration

5.4.2 SME-specific barriers in this category

The obstacles to customer service are organized into three groups. The first is data insufficiency. For a chatbot to work, it needs examples of questions asked by customers, and what constitutes a good answer, and that sort of material exists for some SMEs in e-mails and chats and people's brains rather than in an FAQ. One remedy for that problem has been a proliferation of generative platforms that can produce an initial version of an FAQ from product documentation, and refine it from a handful of seed examples, rather than needing a large corpus of training data [65].

Integration with a client's existing customer relationship management (CRM) or order management system is the second challenge. An SME CRM tool may not have a well-developed API, which will reduce how useful the chatbot might be if it is not able to view an order or update the customer's record [29]. That same review also identifies integration issues as one of the practical hurdles for conversational AI, alongside technology, budget, and data privacy issues [82]. Similarly, for the document automation category, the recommendation is to select tools that can directly connect to an SME's existing customer-facing tools.

The third barrier is brand voice consistency and customer trust. If a chatbot is not aligned with a firm's tone or brand identity, gives out inaccurate information about a product or fails to escalate to a high-priority issue at the right time, then the risks may outweigh the time savings. For instance, customer trust and brand risk are included among the barriers identified in the SME literature [29], which is why a practical remedy that emerges is an approach of a tighter human-in-the-loop: where the chatbot identifies a high-confidence case it answers directly, flags cases of uncertainty, and escalates the higher-stakes questions to a business owner or assigned individual.

5.5 Category 3: Marketing and Sales Optimization

The third category of the taxonomy is Marketing and sales: a group of issues that focus on creating demand, moving a prospect from interest to a transaction, and keeping the customer once he is acquired. This is the category where data from Eurostat has SMEs at a higher level of AI use [77], in part because the tool set (e-commerce stores, email tools, and social media scheduling applications) was already available for use by SMEs [28]. Moreover, it has long been evident that marketing has been a problem that is too big for small firms to deal with in a traditional way.

5.5.1 Definition and scope

Marketing and sales is structurally divided into a low-data half (content generation, campaign optimisation) accessible to most SMEs immediately, and a high-data half (lead scoring, personalisation) that requires first-party customer data many SMEs lack [65, 83]. The category encompasses four distinct workflows. Content creation is the first workflow, which entails generating email marketing messages, social media updates, ad copy, blog posts, and product descriptions. Second is lead scoring and qualification; the ranking and prioritisation of prospects who are most likely to convert into customers. Third is personalisation, in which content, offers or recommendations are tailored to the needs and wants of each individual customer, based on their behaviour or self-declared preferences. And fourth is campaign optimisation, in which alternative versions of content, channel, or timing are chosen based on performance data.

In this category, the two prevailing AI capabilities are generative (used for content creation) and predictive (used for scoring, segmentation, and optimization). Their data requirements could not be more different. The first capability now has low local data requirements; a reasonably capable language model is enough to deliver a good first draft with the help of a few sentences from existing firm communications and some data about the product. Predictive scoring, on the other hand, requires historical customer data, which SMEs without access to a CRM containing several hundred customer data points will not have. The split therefore results in this category being divided into a low-data and a high-data half, with SMEs typically starting with the low-data half.

5.5.2 Why the category matters to SMEs

Eurostat information in Chapter 3 indicated that, at 34.7 percent of AI-using firms overall, marketing and sales is the largest single business case where AI is put to work in enterprise environments [77]. Small and medium-sized enterprises were well-represented in this proportion, partly because the basic tools that AI enhances were already reasonably inexpensive, so implementing AI into them required a fairly marginal investment of additional funds and did not necessitate an all-new investment decision. This same dynamic emerged in the SBA survey, where firms on the smaller end of the size distribution led firms on the larger end in terms of automation usage for marketing, a reverse distribution of leadership for this type of software deployment [38].

As the 2025 New Survey Evidence on Generative AI and the SME Workforce describes, SMEs are now employing generative AI [39, 65]. This publication stands as one of the first studies to examine SME-level language model adoption at scale and reports that content production for marketing communications is one of the most common use cases, as well as the leading use case in areas in which the models interface with customers. The same survey also looks at how the introduction of generative AI is affecting SME skill and labour needs; the latter is especially germane for the marketing and sales category because the types of labour that an SME must supply are dependent on what the language model can take over versus what it must still rely on human labour to do.

The earlier 2021 study of the Digital Transformation of SMEs frames the broader pattern that makes marketing and sales such a productive AI category for small firms [28]. The study notes the economies of scale, transaction cost benefits, and product differentiation that SMEs have achieved through digitization, and that, in particular, they are most catching up to large firms in the area of customer facing platforms. AI builds on this existing structure, which is why the entry barriers are lower in marketing and sales than other categories where the underlying digital base is yet to be established.

Four sub-types of marketing and sales problems recur across the SME population.

The initial sub-category includes content generation. For instance, a marketer in a SME organization who used to dedicate 0.5 days writing a newsletter, 2 days planning a month worth of social media posts, and another day writing the product description of their catalog update would now be able to generate these drafts in a fraction of this time and use the remaining time for reviewing, brand consistency check, and campaign effectiveness measurement. The recent survey on the use of generative AI by the SME workforce covers some examples of this across several OECD economies. The survey finds that SMEs in OECD economies tend to adopt content drafting in particular when starting to use generative AI, and that time saved by doing the same work through the use of generative AI in the SME workplace are significant enough for firm-level effects to be observed [65].

A second sub-type is lead scoring and qualification. A small and medium sized business (SME) in the business-to-business (B2B) space with a few hundred prospects in its sales funnel can make use of a machine learning model to rank prospects by their probability of conversion. Inputs can include size of the lead, level of engagement, how recent the engagement has been, and how the prospect came into the organization in the first place. The scoring does not have to be sophisticated to be valuable; even a simple scoring model that sorts the prospects into three buckets and automatically schedules the top tier for a sales call can be effective at reducing wasted time and effort. The amount of historical data required is fairly small by enterprise standards: low hundreds of leads with conversion outcomes.

A third sub-type concerns the personalization of customer-facing content. An AI-enabled personalization pipeline for customer retention in investment platforms exemplifies the architecture pattern [83]. That pipeline processes data from various sources, like transaction history, portfolio management behaviors, and user interaction logs, to personalize content generation, product recommendations, and user engagement mechanisms. The architectural pattern demonstrated can be easily recognized in the small to medium enterprise (SME) deployment at a lower scale. The underlying building blocks (data ingestion, behavioral modeling, and content selection) remain the same for a vertical SaaS product, where the small and medium-sized business can subscribe to an existing product rather than building the pipeline internally. Additionally, it is observed that traditional rule-based recommendation algorithms fall short in capturing the nuanced, dynamic patterns of investor behaviour and preferences in the current financial market [83], which is the case for most consumer-oriented SMEs as well. Rule-based approaches for customer segmentation are becoming less viable as customers increasingly expect that at least a minimal degree of personalization should apply to their customer experience.

The fourth sub-type is campaign optimisation. For the SME that runs paid advertising on social media or search engines, the relevant AI-augmented functionality now offered natively through those channels includes automated bidding, audience expansion and creative variations. While the SME does not build the underlying models, it determines how to use them and to what extent to entrust them with the optimisation task. This is where, according to survey evidence on SME adoption of generative AI [39], a persistent gap emerges: SMEs have access to AI on platforms but do not systematically run their own campaigns in a way that will generate insights from the results.

Table 5.4. Marketing and sales sub-types classified by data requirement. The low-data half (content generation, campaign optimisation) is accessible to most SMEs immediately; the high-data half (lead scoring, personalisation) requires first-party customer data that many SMEs lack.

Sub-type	**AI capability**	**Data requirement**	**Entry barrier**	**Data half**
Content generation	Generative	Low: product info + brand context suffice	Low	Low-data
Campaign optimisation	Predictive (platform-native)	Low–Medium: platform provides performance data	Low	Low-data
Lead scoring & qualification	Predictive (ML)	High: hundreds of leads with conversion outcomes	Medium	High-data
Personalisation	Predictive + Generative	High: behavioural data, transaction history, interaction logs	High	High-data

In the category of marketing and sales there are less hard barriers than other categories, but two soft barriers are very prominent.

The first is first-party data scarcity. Personalisation and scoring requires first-party customer data, and many SMEs operate via platforms (marketplaces, social commerce, and third-party booking platforms) that control the customer relationship and provide limited access to their data [19]. In such cases, AI in marketing reduces to content generation and optimisation of the platform, and the more advanced use cases related to personalisation are lost.

Second, attribution gets more complicated. With AI-driven campaigns operating across multiple touchpoints, tying revenue or leads to a particular spend becomes harder, not easier. The wider body of work on SME digitalisation [28] concedes that measurement gaps are a persistent issue, and the resulting advice to SMEs is to use the most basic channels-level reporting that can be independently validated as opposed to the most advanced attribution model that the tool can support.

One more challenge, perhaps a bit more nuanced, and one worth highlighting: brand drift. Generative AI content is quick, but content that drifts over time and moves away from the brand's voice is damaging to the trust a brand has developed. The evidence shows that 41.9 percent of SMEs using generative AI report that creativity and innovation skills have become more important since adoption [65], suggesting that the human editorial function has not been eliminated but rather reshaped. Mitigation includes maintaining a style guide that can be fed in to the prompt, reviewing all AI content before publication, and rotating the human reviewer to catch drift.

5.6 Category 4: Knowledge Management and Internal Search

In contrast to the preceding three categories, the fourth deals with "inside out" information flows: employee information search, capturing and transferring new employee know-how, making disparate files searchable, and capturing and making the expertise of senior people available to junior people without constant interruption. This category has also seen the most dramatic technology changes since 2022, with the advent of "retrieval augmented generation" [84], a way of grounding language model responses in a firm's own documents.

5.6.1 Definition and scope

Knowledge management and internal search, as defined in this context, encompasses four major categories of problems. First is internal search that goes beyond the keyword-based search provided by legacy intranet search tools, performing a semantic search of the firm's corpus that matches the question rather than just the words. Second is question answering over company policy, procedures, and products: an employee asks a question and is given an answer with citations based on the firm's documents. Third is knowledge capture, the workflow of turning a senior employee's experience into reusable assets in the form of documented procedure or training a model. And fourth is locating expertise: the "who in the firm knows about X?" question, and the use of communication patterns and document authorship to answer it [29].

The main AI capabilities in this group are those that combine retrieval and generation. Retrieval, typically done with embedding models and vector databases, finds the appropriate documents or passages. Generation, typically done with LLMs, creates an answer, citing or referencing the retrieved content. This combination, called retrieval-augmented generation (RAG), differentiates category 4 from category 1 (where AI ingests inbound documents) and category 3 (where AI produces outbound content).

5.6.2 Why the category matters to SMEs

For small and medium enterprises, this is no novel problem. Chapter 3 described the structural difficulty of fragmented information systems, data scattered around the spreadsheet and email inbox and in random file folders. Chapter 3 also explored this category's root cause, the so-called single source of truth problem. Without a unified data system, there is no one place to find a particular piece of information. Instead, the team must search through multiple tools manually and spend time doing it. The expense of the search, time that small teams do not easily have, is the real cost of missing data. A report on empowering SMEs from data states the problem explicitly, identifying, as some of the key challenges for small and medium enterprises, that there is "lack of data policies and clarity on specific roles and responsibilities," "obstacles to extracting value from the data," and "limited information technology (IT) infrastructure" [85]. All these are limits AI-based knowledge management systems could only partially address, because the constraint is organizational and not only technological.

The literature on industrial use of RAG is peer-reviewed but matured very fast. The 2025 paper on application of retrieval-augmented generation for interactive industrial knowledge management, in Computer in Industry, described deployment of RAG with large language model for interactive access to industrial knowledge [86]. It accumulated citations rapidly within months of publication, an uptake rate that is fast for a peer-reviewed engineering journal paper and indicative of how quickly the technology pattern has matured. The technology deployment described in the paper is industrial, but not specifically SME. The technology pattern described (parse, embed, retrieve, generate with citations) is the same pattern that SMEs will leverage in their SaaS products on top of language model.

In a systematic review, knowledge management is highlighted as an SME activity for which the literature notes both the potential and the barriers [29]. Data quality, integration, and the lack of internal expertise are the same barriers and they, again, have a special relevance in knowledge management because the data to be searched is company specific, so there is no standardisation by any external provider. While in category one document automation benefits from the fact that invoices are structured, in knowledge management the firm has to deal with a history of internally produced documents of varying formats, inconsistent metadata, for which they are all the time the only provider.

5.6.3 Sub-types of problems inside the category

Four sub-types of knowledge management problems recur in SMEs.

The first kind is semantic search within documents. The typical use case is a sales team that needs to locate relevant product specs, technical notes, and previous emails with a customer to answer a customer question. The keyword searching of the existing file and email systems fails, because the customer used a different vocabulary than the source document. A semantic search engine that puts both the question and the source documents into a single semantic space will match the question to the document based on meaning rather than matching words, so in practice the sales team will get an answer in seconds instead of minutes or hours.

The second sub-category is answering questions by referencing source material. In this case the system does not merely provide the relevant documents but also constructs an answer to a given question that draws from multiple documents with references to their source. The RAG pattern allows this type of functionality, while [86] shows the same pattern deployed in an industrial setting on a large scale. For a company the benefit extends beyond the speed of an answer to the auditability as well, as an agent answering a customer from RAG generated response would be able to show the source document if so required. Such a practice is not enforceable with the traditional method of pulling information from memory.

The third subcategory is knowledge expertise location and capture. As an SME scales up, finding out “who knows” can get difficult to figure out, especially when there are 50-250 people in the company and there is no way to ask informally anymore. Tools can help identify expertise through analysis of document authors, code commits (for tech companies), and communication patterns to find possible expertise candidates. Tools could then assist with expert knowledge capture by helping to surface the knowledge being kept in the head of an expert before they leave (a common concern in a small company when a loss of knowledge can be fatal for the entire company). A report on knowledge-enriched agentic AI workflows [87] identifies this as one of the pillars of future AI agents and argues that knowledge graphs act as a shared framework for storing and organizing organizational knowledge, positioning them as part of the foundation for AI agents. The same report notes that using knowledge graphs in AI workflows enhances data quality and consistency by enforcing standardized data models and structures. The use of knowledge graph technology may still be ahead of most SME needs, although it is likely where the state of knowledge practice is heading.

The fourth sub-type of policy and procedures question-answering, in which small businesses have many policies, about data protection, employment, customer terms and conditions, supplier relationships, health and safety, and the like, but employees do not normally read through them all. Questions about them come up in unpredictable ways. By answering them using a RAG-based assistant that sources its answers from the actual policy documents, small firms are helping the small number of staff they employ who act as de facto policy experts, making the policies that actually get read and used in practice.

5.6.4 SME-specific barriers in this category

Three barriers are central to this category and worth taking seriously.

First is the prevalence of document silos and uneven metadata. Many SMEs keep their documents spread across email archives, file sharing and cloud storage systems, instant messaging logs and on individual personal computers. Creating a knowledge management system in this environment would entail either consolidating the documents (slow, expensive and politically fraught) or developing software connectors to each of these sources (which is quicker but harder to maintain) [85]. The report’s suggestion that SMEs take stock of existing IT assets and define their data requirements is a logical first step, though the reality is that most SMEs would have to start somewhere and that somewhere will typically be their shared drive.

The next hurdle is privacy and data protection. When the document corpus contains

Table 5.5. Sub-types of knowledge management and internal search problems, with their primary AI capability, data requirements.

Sub-type	Primary AI capability	Data requirement
Semantic search within documents	Embedding models + vector DB	Medium: internal document corpus must be indexed
Question-answering with source citations	RAG (retrieval + generation)	Medium: curated document set with consistent metadata
Expertise location & knowledge capture	Graph analysis + NLP	High: communication logs, authorship data, org structure
Policy & procedures Q&A	RAG over policy documents	Low–Medium: existing policy PDFs, employment docs

personal data of employees or customers, the 2025 report on AI Privacy Risks and Mitigations for Large Language Models [88, 89] is the go-to document for guidance on privacy and data protection. The guidance was produced as part of the Support Pool of Experts programme. It focuses on privacy risks with LLMs including those in RAG settings and provides patterns to address privacy risks that do not require SMEs to become compliance experts. One such pattern is simple: avoid building a corpus of documents that includes unredacted personal data, or design a RAG pipeline that screens out personal data at the retrieval point.

The third obstacle is the absence of internal IT resources to upkeep the system in the field. RAG needs to be re-indexed when documents change and it degrades unnoticed if a part of the corpus goes stale [90]. What has emerged as a mitigation pattern is to use a SaaS-delivered knowledge management product that takes care of re-indexing on a user's behalf, exposing an administrative console that is as simple for the SME to manage as possible. Chapter 6 discussion of solution patterns will return to this matter.

5.7 Category 5: Decision Support and Analytics

The fifth and final category in the taxonomy is decision support and analytics. It includes the predictive and analytical applications of AI to help SMEs make better operational decisions: What inventory to order? When will demand peak? Which customers have a high risk of churn? Where along the process are anomalies occurring that need to be investigated? This is the deepest category in the taxonomy in that it requires the longest prior investment in data infrastructure, and the most complex combination of capabilities. This is also the category in which the SME gap compared to large firms is among the most pronounced [65, 77], which is why it is at the same time the hardest to recommend but also among the highest in latent value.

5.7.1 Definition and scope

For the purposes of this thesis, the analytics decision support encompasses four primary work flows. The first is demand forecasting, the prediction of the amount of a certain product or service that will be needed in a later period. The second is churn or attrition prediction, the identification of customers/subscribers who are at higher risk of leaving. The third is anomaly detection in operational data, where unusual patterns or trends are detected and flagged by the system as potential fraud, equipment malfunctions, quality issues, or other anomalies of interest. The last is augmented analytics for business intelligence, where a small and medium sized enterprise (SME) owner/manager can simply query their own data using plain natural language and get analytical results without writing a SQL or manually developing a dashboard.

The main AI capabilities present in this category are predictive (ML models for prediction, classification and anomaly detection) and generative (language model interfaces to data) and both of them need the SME to have sufficient quality and volume of structured historical data to fuel the models. It is this data requirement that differentiates this category from the other ones and it is the main reason for its higher barrier to entry.

5.7.2 Why the category matters to SMEs

Research presented in Chapter 3 put decision support and analytics in the category where SMEs are furthest behind large firms. Chapter 3 cited the U.S. SBA survey, which finds small firms about 10 percentage points behind large firms in data analytics applications of AI [38], and Eurostat data, which shows AI in production processes at 19 percent of small firms using AI and 33.5 percent of large firms using AI [77]. Chapter 3 also noted that this is not merely a time gap to be overcome: structural differences in both data availability and analytical capability mean that only SMEs with an ongoing investment over time in their data infrastructure can put decision support AI to work, and SMEs that have made no such investment will not, no matter how cheap the AI technology may become.

The finding from the systematic review is that data availability is the foundational barrier across the SMEs whose barriers to the adoption of AI were investigated [29], and it is particularly tight in the category of decision support. With 27 distinct total barriers being identified in that systematic review, per the protocol of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), "lack of data" was ranked relatively high among the barriers. However, nearly all of the applications to assist decision-making in this area fall apart as soon as this barrier is confronted.

That said, the value proposition is real. An analysis of AI-powered decision making in banking [91] illustrates how predictive capabilities generate significant incremental value when deployed at scale; for SMEs in data-rich sectors (retail, hospitality, professional services), the value calculation follows a proportionally similar logic. The constraint is not the potential but the data readiness.

Four sub-types of decision support problems recur in SMEs.

The first sub-type is forecasting demand. For the SME that carries inventory or provides services that expire quickly, the business model is built upon the ability to forecast what will be required when. Unfortunately, the traditional approach, relying on owner's gut memory or running a simple moving average in a spreadsheet, results in significant missed opportunities. For a concrete example, a 2025 study performed a systematic comparison of four machine learning models for forecasting demand for a food manufacturing SME [92]. The study took data spanning three years for monthly demand across thirteen product categories for one food manufacturer in Nigeria to evaluate LSTM (Long Short-Term Memory), Facebook Prophet, XGBoost and a Gradient Boosting Regressor. In terms of RMSE and MAE (the lower the better) and R^2 (the higher the better) performance scores for testing on unseen data, LSTM was found to consistently outperform the other models. This result was explained by LSTM's ability to capture the complex non-linear and temporal patterns inherent in demand forecasting more effectively than its simpler competitors. Although Prophet was slightly less performant in these scores, it was deemed to provide more interpretability at a cost of a moderate decrease in performance. XGBoost, on the other hand, performed the worst and was not recommended. The point for the SME is not that LSTM was the best, it is that depending on the data, a simple model comparative can be a worthwhile investment.

The second subtype is churn or attrition prediction, which can benefit the businesses with a subscription-based service or recurring revenues (e.g. [29]). SMEs can predict at what risk a customer is to churn and therefore take retention action in advance of the customer churn. These models are relatively inexpensive to build as data requirements are relatively small by the enterprise standard (i.e. a few thousand records at most, past customers and their retention outcomes), and the underlying modeling methods are not particularly challenging [29]. The problem here may be more operational than data analytical, as predicting an at-risk customer is only useful if the company knows what they will do with that customer, i.e. has a retention play.

Anomaly detection is the third subtype and focuses on monitoring operational data. If the SME has a stream of data processing through it such as transaction records, sensor readings and log files then anomaly detection models can be employed to flag anomalies for the SME to investigate. Use cases will be different from industry to industry such as detecting fraudulent transactions in payments businesses, anomaly detection for quality assurance in small manufacturing plants and the monitoring of equipment health in logistics operations. The common thread is that the model does not need to explain why the anomalous activity occurred; it simply needs to provide enough specificity around the anomaly such that an SME analyst can then investigate it.

The fourth sub-type of augmented business intelligence allows the SME owner to interact with their own data via natural language. The existing dashboard model forces either an expensive BI specialist to build and maintain reports or the SME owner to gain a deep technical understanding of the BI tool. The augmented BI model replaces both with a natural language interface to the data, and is included as an area where AI will create productivity gains for the enterprise in the 2026 State of AI in the enterprise report [70]. While this use case is still in the early phases of adoption, the path is clear: the cost of the foundational underlying language models has fallen, integration with common business platforms for SMEs is increasing, and the entry bar is coming down.

Table 5.6. Sub-types of decision support and analytics problems, with their primary AI capability, data requirements.

Sub-type	Primary AI capability	Data requirement
Demand forecasting	Time-series ML (LSTM, Prophet)	High: multi-year transactional history across product categories
Churn / attrition prediction	Classification (ML)	Medium: few thousand customer records with retention outcomes
Anomaly detection	Unsupervised ML	Medium–High: continuous operational data stream (transactions, sensors, logs)
Augmented business intelligence	NL-to-SQL + LLM	High: structured data in ERP/CRM with clean schema

5.7.3 SME-specific barriers in this category

Of all the categories in the taxonomy, Decision support and analytics has the highest residual barrier height, as survey evidence finds that nearly half of SMEs report that employees lack the skills required for AI deployment [65].

The first barrier is historical data. The training data required by predictive models needs to contain outcomes, and if an SME lacks a CRM used steadily over several years, or lacks a transactional system that logs the correct events, the resulting models are likely to be unreliable. Overcoming this obstacle is not a technical challenge; it is an organizational one. If an SME wishes to move into this category, it usually needs to dedicate a sustained period to instilling data discipline before it makes sense to introduce predictive elements.

The second is the lack of a data team. The categories preceding this one do not require analytics personnel, as the processes involved are well defined and the tools are built for non-experts. What decision support requires is for some person in the firm to figure out what the output from the model means, keep the data flowing, and then make decisions based on the predictions. The workaround has been managed analytics services that provide both the technology and the talent, though these require a higher level of investment than the self-service tools available in the preceding categories [36].

The third barrier to value of AI in SMEs is unclear returns on investment. A forecasting model which increases accuracy by 5 percentage points may or may not generate real savings, depending on the SME's real procurement process and whether the model output has been used. To mitigate this, it is necessary to instrument the decision process, and to measure the difference of what is purchased with and without the use of a model. Again, this is rarely done in SMEs and is partly responsible for the difficulty in showcasing value from this category, even when a well performing underlying model has been found. The difficulty of determining the contribution of AI is cited as one of the barriers identified in the systematic review [29], and the difficulty is especially acute in decision support.

5.8 Chapter summary

This chapter has grouped the heterogeneous problems facing SMEs into five AI-ready categories, ordered by the height of the barrier an SME must clear to enter each. The most accessible is document and process automation: the technology is mature, the deployment model is SaaS-first, and the return is measurable in hours recovered. Customer service and engagement follows closely, sharing the same low-integration, plug-and-play character while adding the requirement that the tool respect brand voice and escalate appropriately. Marketing and sales sits in the middle, split between a low-data half that any SME can adopt immediately for content generation and a high-data half whose lead scoring and personalisation depend on first-party customer data that many SMEs do not hold. Knowledge management and internal search rises higher still, because the corpus to be searched is internal, fragmented, and subject to data-protection constraints, so value depends on organisational work as much as on technology. Decision support and analytics stands at the top of the gradient, requiring sustained historical data, integration, and analytical capacity that only firms with a prior investment in data discipline can supply.

Looking across this ascending order, one constraint keeps recurring, though at different magnitudes in each category: skills and knowledge. At lower-barrier categories the skill level need only be sufficient to frame the right question for the tool, identify its errors, and escalate when something goes wrong; at higher-barrier categories, whether the requisite skill and knowledge are present is decisive for model reliability and decision quality. Skills and knowledge are therefore the constraining factor that spans the whole taxonomy, and the single most consequential variable in whether an SME extracts value or stalls.

These five categories form the working structure for the remainder of the thesis. Chapter 6 links each category to its solution patterns and tool families, and Chapter 7 turns the same categories into a staged implementation playbook an SME can follow from problem selection to operation.

Chapter 6

Solution Patterns and Tool Families per Problem Category

6.1 From Problem Categories to Solution Architectures

6.1.1 What this chapter delivers

Building on the five problem classes established earlier, this chapter turns each one into a concrete solution architecture. The analysis proceeds across three levels: solution patterns (the recurring logic for solving each problem class), tool families (the components or products that implement the patterns), and architectural guidance on the trade-offs SMEs face given limited engineering resources and budgets [18, 93]. The chapter closes with the cross-cutting build, buy, or hybrid question and bridges to the implementation playbook in Chapter 7.

6.1.2 Three levels of analysis

These three levels are distinct, not redundant. A pattern is the logical flow that AI follows to solve the problem: a retrieval-augmented generation pattern, for instance, retrieves the relevant company documents, passes them to a language model, and returns a grounded answer. The tools are the components that implement the pattern: in the RAG example, a vector database, embedding models, and orchestration frameworks. The architectural decision for the SME is which combination of pattern and tools fits its data maturity, technical resources, and budget. The OECD's G7 SME adoption paper offers a four-level taxonomy of complexity, embedded, off-the-shelf, customised, and frontier [36]. The first two characterise most SMEs, and the solution patterns in this chapter target that reality.

6.1.3 The build-versus-buy spectrum

The question of solution architecture is whether to build from scratch; or to assemble from open-source and API solutions; or to buy the packaged AI-enabled solution, with the AIaaS model defined to include AI and cloud computing to enable the enterprise use of AI without the on-premise complexity [18]. This has been the focus of particular scrutiny on SMEs, where the aim is to lower the barrier of the entry cost to AI by switching from a capital expenditure to an operating expenditure model [93]. A 2025 survey of 495 enterprise AI leaders showed that 76% of the deployed AI solutions were bought as opposed to built in-house, up from 53% a year before [94]. This tendency towards buying first is likely to be more pronounced in the SME sector, with the engineering resource for the building of a custom solution rarely available [36]. In this chapter, each solution-pattern and class of tools is situated in that context of build versus buy.

6.2 Solutions for Document and Process Automation

Document and process automation, defined in Section 5.3 as the extraction of structured data from unstructured sources, the generation of standard business documents, and the automation of administrative tasks, is the SME category with the lowest remaining barriers and the shortest time to value [29].

6.2.1 Solution patterns

This section groups the solutions for document and process automation into three recurring patterns. The first is OCR-plus-LLM field extraction, where a document image or PDF is ingested by an optical character recognition engine and converted into machine-readable text, and that text is processed by an LLM to identify and categorize fields such as supplier name, invoice number, date, line items, and tax code; the output is validated by business logic and then piped into an accounting or ERP system. This pattern has developed quickly since 2022, combining the high accuracy of OCR and the ability of LLMs to perform contextual identification of fields. A systematic review found that 63.6 percent of enterprise RAG deployments use GPT-based LLMs, and 80.5 percent utilize out-of-the-box retrieval pipelines such as FAISS or Elasticsearch [84].

The second pattern is the intelligent document processing platform, which embeds the OCR-plus-LLM pipeline into an end-to-end service rather than stitching together the extraction, validation, and routing components manually. An IDP platform offers a unified interface to ingest documents, map fields, manage exceptions, and output to targets. For an SME, the value proposition is abstraction of the technical implementation: a firm specifies the document types it must ingest and the fields it wishes to capture, and the IDP platform performs the rest. Subscription-based pricing lowers the adoption barrier compared with capital-expenditure purchases [93]. As only 17 percent of small enterprises are reported to use AI technologies as of 2025 [95], this low adoption base suggests substantial headroom for IDP solutions among businesses that have not yet automated their document processing workflows.

The third pattern is robotic process automation with cognitive augmentation. Whereas traditional RPA is designed to automate routine, rule-based tasks via a software script that interacts with existing applications, AI-enhanced RPA can extend the process automation scope to include exceptions which would normally require human assistance. An IEEE Access review of intelligent document processing and RPA argues that the addition of AI to RPA expands the technology's reach beyond rule-based automation to tasks that involve pattern recognition, classification, and predictive judgement, with measurable gains in accuracy and throughput [96]. It has been noted that AI technologies are generally employed for complex cognitive tasks, whereas RPA is designed for simple, repetitive, rule-based activities, and that combining these two technologies without a clear division of labour can produce sub-optimal outcomes [97]. A case study of an SME that introduced RPA and AI into the order-entry process demonstrates that the combination may be efficient even in a process with low repetition and high customisation [79].

Market studies provide concrete measures of document automation's potential. In 2024, the intelligent document processing market was worth \$2.30 billion, and will grow to \$12.35 billion by 2030, an annual growth rate of 33.1%; the smallest segment, SME, is growing the fastest, although large enterprises still account for the majority of the market [98]. This is because OCR has matured, AI language models have gained sophistication, and subscription pricing makes these tools affordable for the first time to many firms that have not yet made the investment. The cost of automated invoice processing, the primary use case for invoice automation, is very well studied. A survey of 212 accounts payable professionals found an average processing cost per invoice of \$9.40, compared to \$2.78 for best-in-class companies; on average, an invoice takes 9.2 days to process, and only 32.6% are processed without manual intervention [99]. For an SME that receives a few hundred invoices per month, the difference between \$9.40 and \$2.78 would directly improve its operating profit margin, and the gap between a 9.2-day average and the near-real-time turnaround expected from AI would free working capital otherwise tied up in the accounts payable cycle. Contract analysis, which is the second identified sub-type in Chapter 5, is also gaining momentum with the introduction of AI tools. In a 2026 future-ready lawyer survey, of the 810 legal professionals that responded across eleven countries, 92% say they are using at least one AI tool daily, with use-cases including document analysis, contract drafting, and legal research; of those using AI, 62% reported saving between 6% and 20% of their time each week [100]. Although this survey concerns legal professionals rather than SMEs specifically, the same class of contract-analysis tools appears to be moving into lower-cost subscription formats aimed at smaller organisations, an early and not yet independently documented signal that adoption may be extending beyond legal firms.

6.2.2 Tool families and their SME fit

The tool families that enable these patterns may be categorized into three groups. The first family, cloud-native IDP platforms, consists of IDP platforms which are offered as SaaS subscription software. These platforms include a number of pre-trained OCR-and-extraction models for standard business documents like invoices, receipts, and purchase orders, and enable the SME to add custom document types through a configuration screen rather than writing code. The pay-as-you-go pricing model is typical in these instances, with the SME paying on a per-document basis and incurring no upfront investments in infrastructure or development resources [93]. RPA suites with cognitive add-ons represents the second group: legacy RPA platforms that have incorporated AI through third-party integrations with language model vendors or via proprietary development. The main benefit to SMEs is that if a firm already uses these platforms for rule-based automation, the same platform can be used for document understanding, without the need to source a new vendor [96]. LLM APIs used directly for extraction tasks is the third group. In the third group, the SME provides the language model API with text extracted from documents, and the API returns structured data. This model offers the most flexibility but requires more development skill to set-up and maintain. European SMEs are reported to be lagging in all types of AI, including AI used for RPA [36], suggesting that the more low skill IDP platform model is the better choice for most SMEs, compared to the LLM API model.

6.2.3 Architectural guidance: SaaS-first versus component assembly

The architecture recommendation for this group of SMEs is SaaS-first. This is because there is a well-defined integration surface: document automation requires connectors to accounting software, ERP systems, and cloud storage. Reviews of leading IDP products note pre-built connectors to adjacent enterprise systems, such as business-intelligence platforms and major language-model providers, alongside partner-ecosystem integrations with ERP and CRM suites such as SAP, Salesforce, and Microsoft Dynamics [101, 102]; ready-to-use connectivity to the accounting, ERP, and cloud-storage tools an SME already uses is therefore a decisive selection criterion for a SaaS-first deployment. It has been found that 67 per cent of SMEs in the EU are operating at a low level of, or very low level of, digital intensity [103], meaning that most SMEs do not have the in-house IT capability to build and maintain their own extraction pipeline. An exception to the SaaS-first rule would arise if the SME's document types are extremely niche or the system being connected lacks a standard API. In those situations, a component-based approach using the LLM API for extraction, with a custom connector for routing, may be required. Even then, the recommendation is to source the extraction piece as a product rather than building in-house. A systematic review confirms that process automation is the area of AI where the gap in adoption between large and small enterprises is the smallest [29], which suggests that the ecosystem of tools is sufficiently mature for consumption by SMEs. An SME considering document automation should therefore begin by checking whether connectors exist for its existing accounting or ERP system before evaluating specific products.

6.2.4 Summary and key decision points

The three most important decision points for an SME to ask when considering document automation solutions are: (1) Do connectors exist for the firm's accounting or ERP system? (Answering yes suggests that SaaS-first is a viable solution) (2) Are the document types standard (e.g., invoice, receipt) or specialised? (Answering standard suggests that pre-trained models would work) (3) Is an RPA platform already in use? (An affirmative answer may mean that cognitive augmentation is the least disruptive path to adopt automation).

Table 6.1. Solution patterns for document and process automation. SaaS-first is the recommended default for SMEs.

Solution pattern	Tool family	Key trade-off
OCR + LLM extraction	Cloud IDP platforms	Accuracy vs. document diversity
Intelligent document processing	End-to-end IDP SaaS	Vendor lock-in vs. ease of use
RPA + cognitive automation	RPA suites with AI add-ons	Existing RPA required
Direct LLM API extraction	LLM API providers	Flexibility vs. maintenance cost

6.3 Solutions for Customer Service and Engagement

Customer service and engagement, as defined in Section 5.4, involves the automation of customer interactions via chatbots, voicebots, and ticket classifiers. Chapter 5 identified customer service as one of the most plug-and-play categories for SMEs, and an estimated 37.3 per cent of SMEs using generative AI use it for customer service [65].

6.3.1 Solution patterns

This section groups the solutions for customer service into three recurring patterns. The first is chatbot-as-a-service. Here, an SME uses a web-based chatbot provided by a platform that also supplies both the conversational AI model and the user interface. In a case study, a micro-enterprise is shown to have implemented a hybrid architecture pairing rule-based conversation flows with natural language processing for intent discovery [104]. The system relied on predefined scenario trees for answering common questions, and NLP to classify customer messages and channel them to the relevant response set. The case study argues that even within a micro-enterprise context, AI-powered tools can meaningfully improve operational efficiency. Much evidence points to increased productivity as a result of using AI in customer service. In a study of 5,179 customer support agents for a Fortune 500 company, the usage of an AI tool was found to increase the number of customer issues resolved per hour by 13.8 percent. This effect was strongest for junior agents, who experienced a 35 percent productivity boost; they also experienced an 8.6 percent reduction in attrition rates [15]. Based on current trends, AI is estimated to increase productivity in the area of customer service by 30 to 45 percent [14]. The second is voice AI platforms. These systems use speech-to-text conversion, natural language model processing, and text-to-speech generation to take calls. An ACL conference review on speech language models reports on this kind of end-to-end conversational model in real-time spoken conversation [105]. Small businesses can use voice AI to overcome one capacity problem: A small business cannot maintain a full-time telephone operator but can still deploy a voice AI to take routine phone calls, book appointments, and transfer issues to a human operator during business hours. A systematic analysis of 50 studies concludes that sophisticated chatbots and voice AI provide 24/7 service, a capability otherwise restricted to larger businesses [106]. The third is ticket triage and classification. In this configuration, incoming customer service questions are classified and assigned automatically by an AI to an appropriate support agent. A ticket classification workflow using BERT and TF-IDF with SVM techniques has been reported on a sample of 4,243 chat-based user queries with classification accuracy capable of providing measurable gains to customer support operations [107]. For SMEs who make use of help desk systems, ticket triage is usually sold as an optional extension to the base package, so it is among the most frictionless AI integrations available.

The economic impact of conversational AI in customer service is substantial. Conversational AI deployments within contact centres were projected to reduce agent labour costs by USD 80 billion by 2026, with labour representing up to 95 percent of contact centre costs and the integration cost for a conversational AI agent ranging from USD 1,000 to USD 1,500 [108]. Looking further ahead, agentic AI is predicted to autonomously resolve 80 percent of common customer service issues without human intervention by 2029, resulting in a 30 percent reduction in operational costs [109]. According to a survey of 6,500 service professionals, 30 percent of customer service cases were resolved by AI in 2025, and AI is expected to handle half of all cases by 2027. Service representatives using AI spend 20 percent less time on routine cases, freeing an estimated four hours per week [110].

The voice AI segment is growing rapidly. The AI voice agents market is valued at USD 2.54 billion in 2025, projected to reach USD 35.24 billion by 2033 at a compound annual growth rate of 39.0 percent, with customer support automation accounting for 44.2 percent of the revenue share and inbound voice agents representing 52.1 percent of the market [111].

6.3.2 Tool families: component-based stacks versus vertical SaaS

Customer service AI tools can be thought of in two different types of product. One is component-based stack, where an SME puts a chatbot on its website, sends messages from the chatbot to an LLM API, then connects the chatbot's answers into its own CRM system. While this strategy provides full control over how the bot conducts the conversation and how its data is handled, it requires some degree of technical expertise to implement and maintain. The integration with an SME's existing CRM or order management systems has been identified as a hurdle to implementing conversational AI [82]. The second is vertical SaaS, an all-in-one AI system, which integrates chatbot and phone answering and ticketing features into a single package. This is the preferred baseline choice for SMEs that do not have an IT team to support it. Customer service automation is considered one of the most appropriate ways for a small business to start on the journey towards digitalisation, and the vertical SaaS business model is in keeping with this since no custom integrations are required [80].

6.3.3 Architectural guidance

Three architectural aspects for SME implementations of customer service AI are important to consider: First, escalation design. All customer service AI needs to include some logic regarding whether and when a conversation should be passed to a human. The barriers of customer trust and brand risk that recur in the SME literature [29] are the primary reasons to include a well-designed escalation path. In practice this could be resolved with some form of a confidence based escalation, where the user is transferred to a human if the AI's confidence level in an answer is below a certain threshold. Second, CRM integration. A customer service chatbot that cannot check order status or update a customer record is of very little use. The cost of this integration may vary depending on whether the underlying CRM has a well developed API available or not. Third, brand voice consistency. An estimated 41.9 percent of SMEs using generative AI say that creativity and innovation skills are now considered more important after the AI tool was adopted [65], so while the editor's role has not been completely removed it has changed. A well written style guide that is incorporated into the LLM prompt is one way of addressing brand drift. The privacy-preserving patterns discussed in Section 6.5.3 are equally relevant here, given that customer interaction data typically constitutes personal data under GDPR [88].

6.3.4 Summary and key decision points

Table 6.2. Solution patterns for customer service and engagement.

Solution pattern	Tool family	Key trade-off
Chatbot-as-a-service	Low-code chatbot platforms	Speed vs. customisation
Voice AI platform	STT + LLM + TTS SaaS	24/7 coverage vs. integration cost
Ticket triage	Helpdesk add-on AI	Low friction vs. limited scope

The key decision points for customer service are: (1) whether the firm's CRM has a well-developed API, which determines the viability of a component-based chatbot; (2) whether 24/7 telephone coverage is a business requirement, which determines whether voice AI is needed; and (3) whether the existing helpdesk platform supports AI add-ons, which determines the ticket triage path.

6.4 Solutions for Marketing and Sales Optimization

Marketing and Sales as defined in section 5.5 involves content generation, lead scoring, personalisation and campaign optimisation. Chapter 5 mentioned that this was a category of activities where the EU data shows the largest AI adoption in the SME sector, where 34.7 percent of AI-using enterprises apply AI to marketing or sales [77], as well as where 49.7 percent of businesses using generative AI are reported to apply it for marketing content [65]. One notable fact about this category is that its main implementation pattern for SMEs is not by buying an AI product, but instead by acquiring AI functionalities bundled into existing marketing platforms [112].

6.4.1 Solution patterns

The main implementation pattern is bundled AI within existing marketing platforms, as most SMEs access AI capabilities through functionality already embedded in their email marketing, social media management, or advertising tools rather than by purchasing standalone solutions. A systematic literature review of 97 papers on MarTech adoption in the SME context confirms that the CRM tool forms the main pillar of the marketing technology stack, widely adopted for its accessibility and strategic benefits, while the integration of generative AI is still in an early stage of development, constrained by low data quality and data maturity in SMEs [112].

The second pattern is standalone content generation, where the SME acquires a general-purpose generative AI tool to produce marketing and sales content for website copy, social media posts, blog posts, and product descriptions. A randomised controlled trial with 453 participants showed that having a generative AI tool available produced a 40 percent time reduction in completing a professional writing task and an 18 percent quality improvement [16]. This pattern requires the least data and the least integration, working as a standalone productivity tool for the marketing function.

The third pattern is CRM-embedded lead scoring and personalisation, which uses AI at the higher data end of marketing, where first-party customer data is required. Machine learning has been shown to be usable for lead prioritisation in a business-to-business setting: an analysis of 23,154 records with 15 classification algorithms correctly classified 98.39 percent of records with a Gradient Boosting Classifier [113]. SME-scale offerings in this space are typically available through CRM platforms that integrate lead scoring into the existing sales process [112]. A pipeline has been described that ingests transaction history, portfolio management behaviour, and user interactions to power a personalised content creation and product recommendation engine [83].

6.4.2 Tool families

There are four types of tools for small businesses in this marketing and sales category. The first are the email, social, and advertising tools that small businesses are already using, which increasingly offer AI-driven content suggestions, subject line suggestions, and audience segmentation features, often at no additional cost [65], marketing platform AI features. The second tool category standalone generative tools is content creation. The third category CRM predictive features is lead scoring and churn risk. The fourth e-commerce platform personalisation engines is product recommendations based on browsing and purchase history. In a survey of 250 small and medium-sized enterprises, the use of marketing AI was found to result in a 42.5% improvement in sustainable business profitability resulting from increased customer engagement (by 50%) and improved decision making (by 76%) due to enhanced data [114].

6.4.3 The bundled-versus-standalone decision

The default advice for small businesses is to use what is provided by the platforms. The marginal cost is effectively zero, as the tool is already integrated and no additional training is needed [18]. The framework of platform-dependent entrepreneurs is relevant here: many small businesses are dependent on platforms that own the customer relationship and have very little access to customer data [19]. In this context, the AI features in platform-based tools are not only convenient but could be the only options available to small businesses because the platform owns the customer data for personalisation. The only exception is the small business that has developed its own first-party data outside of the platform (for example, using its own customer relationship management system) that requires capabilities that cannot be provided by the platform. In such cases, a stand-alone lead-scoring tool or a personalised recommendation engine could make sense, though firms with such independent first-party data capabilities are reported to remain a minority [65].

The revenue impact of AI personalisation has also been measured. Businesses with higher growth rates derive 40 percent more of their total revenue from personalised experiences than their slower-growing peers; personalisation typically boosts revenue by 10 to 15 percent, though outcomes can vary between 5 to 25 percent depending on the industry and a company's specific ability to execute, and on the consumer side, 71 percent of people anticipate companies will deliver tailored interactions, while 76 percent report frustration when they do not receive them [115]. For SMEs operating on e-commerce platforms, personalisation has therefore shifted from a differentiator to a baseline expectation.

The use of generative AI for content generation is now the norm in the marketing world. A 2026 state-of-marketing survey, which polled over 1,500 global marketers, indicates that 80 percent of marketers already use AI to create content, while 94 percent plan to expand their usage in 2026. Given that small businesses are 23 percent more likely than average to see a return on blog posts [116], there is a strong case that AI-powered content generation offers smaller firms outsized value. Yet, 84 percent of marketers admit to running generic campaigns despite the availability of personalisation tools [110], suggesting the technology is advancing faster than most companies can leverage it.

6.4.4 Summary and key decision points

The key decision points for marketing are: (1) whether the firm already subscribes to a marketing platform with embedded AI features, which determines whether bundled AI is the default; (2) whether the firm owns sufficient first-party customer data outside the platform, which determines whether standalone lead scoring is justified; and (3) whether content production is a bottleneck, which determines the value of standalone generative tools.

Table 6.3. Solution patterns for marketing and sales optimisation.

Solution pattern	Tool family	Key trade-off
Bundled platform AI	Marketing SaaS features	Convenience vs. platform lock-in
Standalone content gen	Generative AI tools	Productivity vs. editorial oversight
CRM-embedded scoring	CRM predictive features	Accuracy vs. data requirements
Personalisation engines	E-commerce platform AI	Revenue uplift vs. first-party data

6.5 Solutions for Knowledge Management and Internal Search

Knowledge management refers to the ability to make the documents within an organisation searchable, queryable, and actionable. Section 5.6 provides a more detailed discussion of knowledge management. As noted in Chapter 5, the knowledge management category has seen significant technological advances since 2022 with the rise of retrieval-augmented generation. A systematic literature review of 63 primary studies confirms that enterprise use of RAG is still at an experimental stage, with fewer than 15 percent of studies addressing the real-time integration challenges required for production-scale deployment [84].

6.5.1 Solution patterns

This section groups the solutions for knowledge management into four recurring patterns. The first is embedding-based semantic search: company documents are uploaded into the system as vector embeddings, which capture their semantic content. A user query is embedded in the same way, and the most relevant documents are returned based on vector similarity. The underlying infrastructure has been classified into three categories: native systems designed around vector management (such as Milvus and Vearch), extended systems that add vector capabilities on top of existing databases, and search engine libraries (such as FAISS and Elasticsearch) [117]. The vector database market was worth \$2.65 billion in 2025 and is projected to reach \$8.95 billion by 2030, with a compound annual growth rate of 27.5 percent [118].

The second pattern is the retrieval-augmented generation pipeline. RAG extends semantic search by passing the retrieved documents to a language model, which produces a natural-language answer grounded in the source information. A practical RAG system for industrial knowledge management has been presented that integrates BM25 keyword search with embedding-based retrieval and a reranker, attaining a mean reciprocal rank of 88 percent, recall of 85 percent, and mean average precision of 75 percent in a technical service [119]. For an SME, the main benefit of RAG is that it avoids training an original model: the language model is consumed via API, and the firm's company knowledge enters the system through the retrieval layer rather than through fine-tuning [84]. The risk of hallucinated responses is a known concern. Chatbots without RAG have been reported to have hallucination rates of 37 percent (GPT-4) and 40 percent (GPT-3.5), whereas RAG with curated knowledge reduced hallucination to 0 percent for GPT-4 and 3 to 6 percent for GPT-3.5 [120]. The MEGA-RAG framework reports a hallucination-rate reduction of more than 40 percent compared with baseline models [121]. The retrieval layer is therefore not an optional add-on but the key enabler of factual consistency [120, 121].

The third pattern is the all-in-one knowledge management platform. These are SaaS products that offer embedding, retrieval, generation, and an administration console as part of a single subscription. For the SME that cannot easily stitch the layers of a RAG pipeline together, an all-in-one platform should be the default choice. A systematic literature review observes that SMEs typically face constraints on digital infrastructure, technical competence, and budget, hindering the adoption of AI-based knowledge management technologies; the all-in-one platform design overcomes this limitation by concealing the technical details under a simple administration UI [122].

The fourth pattern, knowledge graph enrichment, looks to future developments. Knowledge graphs are noted to improve data quality and consistency in enterprise AI by enforcing uniform data models and structures [87]. Knowledge graphs may not yet be applicable to every SME, but they represent the next step for enterprise knowledge management.

The productivity gap among knowledge workers has been widely measured, providing a compelling business case for knowledge management AI. These workers are estimated to spend nearly 20 percent of the workweek searching for and gathering information. By using technology to improve internal search and collaboration, knowledge workers could be made 20 to 25 percent more productive, unlocking $900 billion to $1.3 trillion in annual value across four sectors [123]. For an SME with 20 knowledge workers, a conservative 10 percent boost in productivity through better internal search amounts to two additional full-time workers at no extra cost.

The enterprise search market meets this need, currently worth USD 7.47 billion in 2026 and forecast to reach USD 11.66 billion by 2031 at a compound annual growth rate of 9.31 percent, with conversational and natural-language-processing-based search its fastest-growing segment, consistent with the RAG pattern noted above [124]. As the industry moves from searching based on keyword retrieval to semantic and conversational search, this represents a major technological inflection point that makes knowledge management AI feasible for SMEs: the technology now understands natural-language questions and returns synthesised answers, without requiring users to be particularly precise with their queries.

6.5.2 Tool families

The set of tools for knowledge management is organised into four layers. The first layer is embedding models, covering both open-weight models for local deployment and API-hosted models from a major provider. The second layer is vector databases, differentiated between native, extended, and search engine types [117]. The third layer is LLM APIs, which includes the models that provide generation for the RAG pipeline. The fourth layer is integrated KM platforms, covering platforms that offer all of the above in a single package. Flowise AI has been used to create an example RAG setup that takes low-code methods: instead of individual pieces of code, the platform provides a visual workflow builder with configurable graphical pieces, which can be assembled without programming to form a Conversational Retrieval QA Chain [49].

6.5.3 Privacy-preserving options

If the document corpus stores personal data on the employees or customers of the firm, the 2025 report on AI Privacy Risks and Mitigations for Large Language Models offers authoritative guidance [88]. A systematic review of AI guidelines for SMEs confirms the requirement for sector-specific guidelines as there is a wide variation in preparedness across sectors [89]. Three privacy-preserving patterns include: storing and processing embeddings and performing retrieval on premises (most restrictive), choosing a cloud vendor that is allowed by data-residency laws, or designing the corpus from the beginning to omit data that is identifiable as belonging to individuals (so such data need never be ingested).

6.5.4 Architectural guidance: corpus maintenance

The least discussed and most under-acknowledged architectural challenge in knowledge management is corpus maintenance. Outdated content in the corpus has been found to significantly degrade the performance of RAG systems in two ways: it severely degrades response accuracy because it distracts the model from accessing the information that the model should use as an answer, and it can cause the model to hallucinate answers that are potentially damaging even if better-quality information exists [90]. A RAG system that is not regularly re-indexed will therefore degrade silently, with answers becoming wrong without any error being raised. The most reliable mitigation for SMEs is a SaaS knowledge-management product that handles re-indexing automatically through a simple administrative interface. It has been emphasised that the challenge is not just technical, it is organizational: many SMEs still maintain their documents dispersed between email archives, file sharing and cloud-storage systems, and instant-messaging logs [85].

6.5.5 Summary and key decision points

The key decision points for knowledge management are: (1) whether the firm's documents are consolidated in a single repository or dispersed across multiple systems, which determines the indexing challenge; (2) whether the document corpus contains personal data, which determines whether privacy-preserving deployment is required; and (3) whether the firm has internal capacity to maintain the corpus over time, which determines whether an all-in-one SaaS or a managed service is preferable. The privacy-preserving patterns discussed in Section 6.5.3 apply equally to customer service and marketing use cases where customer interaction data constitutes personal data under GDPR.

Table 6.4. Solution patterns for knowledge management and internal search.

Solution pattern	Tool family	Key trade-off
Embedding-based search	Vector DBs + embedding APIs	Precision vs. setup complexity
RAG pipeline	Vector DB + LLM API + orchestrator	Accuracy vs. maintenance burden
All-in-one KM platform	Integrated KM SaaS	Ease of use vs. customisation
Knowledge graph enrichment	Graph DB + LLM + ontology	Future value vs. current complexity

6.6 Solutions for Decision Support and Analytics

Decision support, defined in Section 5.7, includes demand forecasting, churn prediction, anomaly detection, and augmented business intelligence. Chapter 5 identified this as the category with the highest barriers and the largest disparity between SMEs and large firms: only 17.4 percent of small firms use AI compared with 52 percent of large firms [36]. The solution patterns below bring decision support within reach of SMEs despite these barriers.

6.6.1 Solution patterns

This section groups the solutions for decision support into four recurring patterns. The first is vertical SaaS forecasting. In this pattern, a piece of domain-specific software delivers a forecasting model that is trained on the data from the firm's own transactions and provides results through the same interface that the firm uses for its daily work. This approach has been applied to inventory management for a food manufacturing SME, where an LSTM model was found to perform better than Prophet, XGBoost and the Gradient Boosting Regressor on the RMSE, MAE and R-squared performance metrics using unseen data [92]. The main advantage for SMEs is that this is integrated into the operational system: the firm neither needs a data scientist, nor a separate analytics infrastructure. The second pattern is managed analytics services. A third party provides both the software and the analytical skills to deliver forecasts, dashboards and recommendations on the SME's behalf. This approach is noted to have a greater cost than self-service alternatives [36], but the skills barrier is the binding constraint identified in Chapter 5 for this category. The third pattern is natural-language business intelligence. By 2027, 75 percent of new analytics content is predicted to be contextualised for intelligent applications via generative AI, while over 50 percent of analytics leaders surveyed have already deployed AI tools to deliver automated insights and natural language queries [125]. For SMEs, natural language BI democratizes data analytics: a firm owner can pose questions using plain language and get back charts, tables, and explanation without needing to write SQL or build a dashboard. A conversational BI system that reduces dependency on

IT teams and speeds up the generation of insights has been described [126]. Further, a natural experiment on 370,000 sellers on a leading e-commerce platform shows that sellers that adopt an analytics product enjoy a 14.4 percent uplift in the number of transactions [127]. The fourth pattern is ERP/CRM-embedded predictive features. Embedded AI is projected to accelerate financial close by cloud ERP by 30 percent by 2028. Cloud ERP spend will be 62 percent AI-enablement by 2027, up from 14 percent in 2024 [55].

6.6.2 Tool families

The four tool families for decision support closely mirror the four patterns. Vertical SaaS forecasting products combine a transaction system with an embedded predictive engine. Managed ML services offer models and talent on subscription. BI tools with natural-language interfaces, with generative AI at their centre, are among the most rapidly expanding, reflecting the augmented-analytics trend noted in Section 6.6.1. And ERP/CRM-embedded prediction is the longest-term investment because the firm's selection of an enterprise platform defines the predictive capabilities available. A study of churn prediction in a Brazilian SaaS company confirms what predictive accuracy is feasible for decision support: a Decision Tree model with 31 features and a dataset of 4,911 records achieved an accuracy of 91.25 percent, showing that a predictive accuracy that is high enough for decision purposes does not require a huge dataset [128]. Finally, a study of 11,429 European SMEs finds that AI-adopting SMEs have a 15.1 percentage point higher likelihood of having revenues growth of over 30 percent [66], which provides empirical evidence that the decision to invest in analytics tools pays off.

6.6.3 Architectural guidance: build versus subscribe

The default recommendation to an SME is to subscribe. Building a custom predictive model requires sufficient historical data of adequate quality, a pipeline to provide the model with fresh data, and personnel who can interpret the output. A systematic review identifies the inability to determine the contribution of AI in a given organisation as a persistent barrier [29]; the subscription model addresses this with predictable cost, a demonstrated value proposition (measurable through A/B testing), and an exit option if results do not materialise. The outlier is the SME that is data-rich by business sector (retail, hospitality, professional services), has invested in data infrastructure, and has analytical talent available in-house or on contract. In such cases a custom model can yield higher accuracy than a standard platform, particularly in demand prediction where local conditions diverge from the distribution on which the platform model was trained [92].

Three independent surveys confirm that organisational AI maturity in decision support remains low. A survey of 2,018 SMEs across twelve countries finds that 76 percent of AI-using SMEs are classified as AI Novices (using basic tools for discrete tasks), with only 8.6 percent reaching the Explorer or Champion levels of extensive integration [129]; Figure 6.1 illustrates this distribution. A further analysis adds a size gradient: while only 18 percent of US companies had adopted AI by the end of 2025, those companies employ 78 percent of the US labour force, confirming that adoption is concentrated among larger firms [130]. A 2025 global survey reports a similar pilot-to-production gap: 88 percent of companies use AI in at least one function, but only 39 percent report measurable EBIT impact, and only 29 percent of firms with revenues below US$100 million have entered the scaling phase, compared with nearly half of those above US$5 billion [43]. For SMEs targeting decision support, the implication is clear: the subscription path is not a fallback but the realistic default, while custom builds are reserved for the data-rich exceptions.

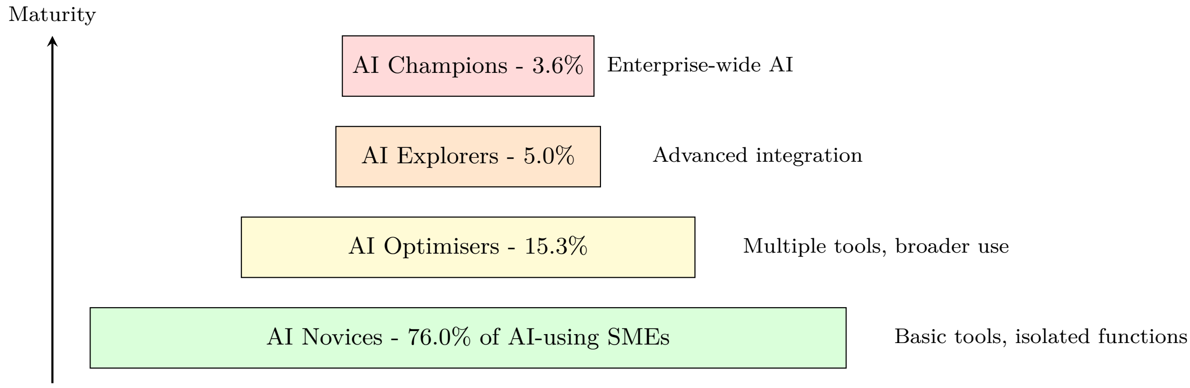


Figure 6.1. SME AI adoption maturity distribution. Data from the OECD D4SME 2025 survey of 2,018 SMEs across twelve countries [129]. The pyramid shape illustrates that the vast majority of AI-using SMEs remain at the novice level, with fewer than 9 percent progressing to advanced or champion status.

6.6.4 Summary and key decision points

Table 6.5. Solution patterns for decision support and analytics.

Solution pattern	**Tool family**	**Key trade-off**
Vertical SaaS forecasting	Domain-specific SaaS	Integration vs. customisation
Managed analytics	ML-as-a-service + talent	Cost vs. skills gap
Natural-language BI	BI tools with NL interface	Accessibility vs. data readiness
ERP/CRM-embedded predictive	Enterprise platform AI	Vendor choice vs. embedded value

The key decision points for decision support are: (1) whether the firm has at least two years of clean historical data in its core operational system, which determines the feasibility of predictive modelling; (2) whether analytical talent is available internally or must be procured as a managed service; and (3) whether a natural-language BI interface would provide sufficient value without a custom model, which is often the case for firms that need insight access more than prediction accuracy.

6.7 Cross-Cutting Analysis: Build versus Buy, SaaS versus Custom

6.7.1 The architectural decision framework

A four-level taxonomy of AI use complexity gives SMEs a useful decision framework for choosing whether to build or buy: embedded (AI built into the tool the firm already uses); off-the-shelf (publicly available AI tools used without integration); customised (models trained on proprietary data); and frontier (cutting-edge models requiring significant in-house capability) [36]. The evidence surveyed in this chapter suggests that the majority of SME AI use cases can be well satisfied with the first two levels in this taxonomy. A 2024 enterprise AI survey reports that 76% of enterprise AI solutions in 2025 were bought rather than built internally (in 2024 it was 53%), and that total enterprise AI investment in the 12 months before that survey reached $37 billion, up 3.2 times from 2024 [94]. Given the scarcity of internal engineering capacity, the buy share among SMEs is likely even higher. For each of the solution pattern categories discussed in the preceding chapters, the following default positions emerge along the build versus buy scale. Document automation and customer service: buy (SaaS platforms for automation and conversational service). Marketing: bundled (tools with AI functionality are bundled into other tools firms already pay for and use). Knowledge management: buy for all-in-one platforms for SMEs, assemble for firms with the technical resources to do so. Decision support: subscribe to vertical SaaS providers and managed service offerings. Firms that have unusual proprietary data conditions, are constrained by particular regulations, and/or have in-house technical teams could consider building in some cases, but the general stance should be to use off-the-shelf solutions when feasible [36, 94].

6.7.2 The agentic AI shift

This section examines the architectural implications of the shift toward agentic AI. The move from single-turn conversational agents toward multi-step agentic AI conversational agents for workflows will change the solution design patterns in most categories. It is predicted that 40% of enterprise applications will have task specific AI agents by 2026, up from under 5% in 2025 [131]. It has been argued that scaling agentic AI requires fully integrated platforms capable of managing enterprise data and supporting the lifecycle of AI applications from build to deployment. The practical implication for SMEs is not to build agentic systems today, but to avoid architectural choices that foreclose the option in the future [72]: prefer platforms that have APIs exposed for programmatic access, event-driven integration patterns, and data stores where data can be moved out in standardized, structured format. It is reported that 95 percent of organisations with investments in generative AI are extracting no measurable return, as they struggle to move out of experimentation mode [132]; SMEs should not replicate this pattern in a few years by being the enterprises that had invested in frontier generative AI capabilities before having solved the use cases for the simpler solution patterns.

6.7.3 Loose coupling and the managed-service default

For each of the five solution pattern categories, the key architectural principle loose coupling is: components can be swapped without affecting the other components. The AI as a service principle captures such a design: the SME consumes AI capabilities through an API, and if a superior model becomes available from a different provider, switching is straightforward [18]. Loose coupling translates into the managed-service default for SMEs. The SME can delegate integration to a platform provider, and instead of having to integrate on the infrastructure level, they can take it as a service. AI adoption across OECD firms has been found to have more than doubled, rising from 5.6 percent to 14 percent between 2020 and 2024, with firm-level productivity premia at least 4 percent and, in some cases, upwards of 15 percent [36].

A 2024 survey of 753 IT professionals found that 89 percent of organisations have adopted a multi-cloud approach, with 73 percent having hybrid cloud implementations; for AI specifically, 41 percent of organisations are currently using machine learning and AI in the cloud, while another 32 percent are experimenting [133]. The gap between general cloud adoption (89 percent) and AI-specific cloud workloads (41 percent) suggests that many firms have the infrastructure in place but have not yet extended it to AI use cases.

Figure 6.2 visualises the default position for each category.

The open-source share of enterprise AI deployments is declining. Enterprise use of open-source LLM APIs is reported to have fallen from 19 percent in 2024 to 11 percent in 2025 [94]. This trend runs counter to the expectation that open-weight models would erode proprietary market share. For SMEs, the practical adoption trajectory favours proprietary APIs and managed services.

6.7.4 Vendor lock-in risks and mitigation

Managed-service convenience, of course, entails a risk of vendor lock-in. It is documented that three cloud providers comprise 68 percent of global cloud-computing market share [17], and the power imbalance between platform and provider is noted to be inherent to the economics and tech architecture of platforms [19]. There are three possible responses for SMEs. The first is data portability: retain ownership of and access to data in a format that can be ported to a competitor. The second is API standardisation: select platforms that use open standards rather than proprietary interfaces. The third is hybrid deployment: run on open-weight models deployed locally for privacy-sensitive or latency-critical applications, while keeping general applications on the cloud API. This hybrid configuration is possible for firms with sufficient technical know-how, though 67 percent of SMEs are reported to exhibit low levels of digital intensity [103], suggesting it may not be a viable option for most firms just yet.

For each of the five categories, the default position is to buy or bundle rather than build: document automation and customer service call for packaged SaaS applications; marketing capabilities arrive bundled within software firms already use; and decision

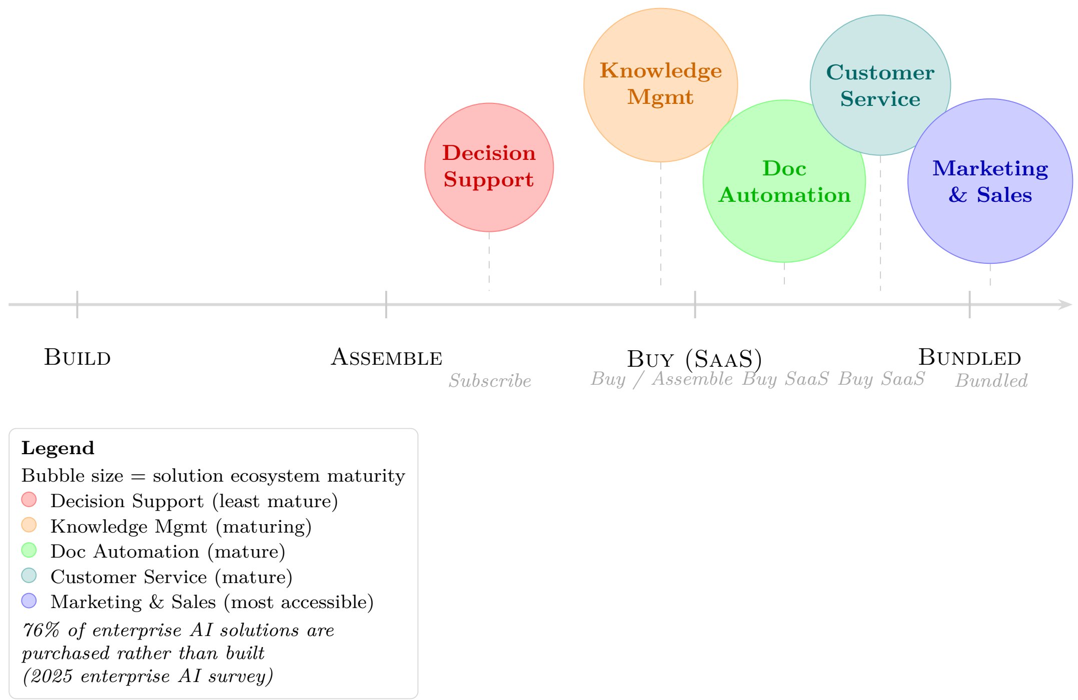


Figure 6.2. Default position of each problem category on the build-versus-buy spectrum. Bubble size reflects the relative maturity of the solution ecosystem: larger bubbles indicate more packaged, SME-ready options. Categories toward the right are consumed through existing subscriptions; those toward the left require more customisation [36, 94].

support is most realistically reached through subscription. Custom-built systems are appropriate only in special cases: non-standard proprietary data, binding regulatory constraints, or firms with genuine in-house engineering capacity. The architectural question is rarely whether to adopt AI at all, but how much integration effort a given firm must absorb. Chapter 7 translates these patterns and trade-offs into a staged implementation playbook that an SME can follow, from problem selection to production operation.

Table 6.6. Summary of the principal sources informing this chapter, by evidence type.

Source (author/year)	Type	What it supports
Lins et al. (2021) [18]	Article	AIaaS model; loose-coupling and consumption of AI via API
Griesch et al. (2024) [93]	Peer-reviewed	AIaaS lowering the cost barrier for SMEs (capex to opex)
Oldemeyer et al. (2024) [29]	Peer-reviewed	SME adoption gap smallest in process automation; recurring barriers
Karakurt & Akbulut (2025) [84]	Peer-reviewed	RAG and IDP deployment patterns; maturity of retrieval pipelines
Brynjolfsson et al. (2025) [15]	Peer-reviewed	Customer-service productivity gains from AI assistance
Noy & Zhang (2023) [16]	Peer-reviewed	Generative AI productivity and quality gains in writing tasks
Pan et al. (2024) [117]	Peer-reviewed	Taxonomy of vector database systems for semantic search
Nishisako et al. (2025) [120]	Peer-reviewed	Hallucination reduction through retrieval grounding
Cutolo & Kenney (2021) [19]	Peer-reviewed	Platform dependence and vendor power asymmetry
OECD G7 paper (2025) [36]	Institutional	Four-level complexity taxonomy; build-versus-buy framing
OECD D4SME (2025) [129]	Institutional	SME AI maturity distribution across twelve countries
Eurofound & Cedefop (2025) [103]	Institutional	Low digital intensity across most EU SMEs
Menlo Ventures (2025) [94]	Consulting	Buy-over-build share of enterprise AI deployments
McKinsey (2021) [115]	Consulting	Revenue impact of AI-driven personalisation
Grand View Research (2025) [98]	Consulting	Market sizing and growth of intelligent document processing

Chapter 7

Implementation Guidelines: presentation of the Playbook

7.1 From Categories and Solutions to a Sequenced Path

This chapter turns the preceding analysis into an operational sequence: a single adoption path an SME can follow from the choice of a first use case to a governed, operating deployment, organised as six sequential, gated phases. Where Chapters 5 and 6 asked which problems AI can address and which tools fit them, this chapter asks in what order a firm should act, and what must hold true at each step before the next begins.

The path is depicted as consisting of six phases. Phase 1, called problem selection, is when the firm identifies a particular use case to address first. Phase 2, readiness assessment, is when the firm undertakes a self-evaluation of its readiness against the needs of the identified use case. Phase 3, technology selection, is when the firm decides how and to what extent to build versus buy and reasons about the full cost of the choice. Phase 4, pilot, is when the firm undertakes a bounded first deployment with an explicit, predefined criteria for what success looks like. Phase 5, integration and change management, is when the firm incorporates the validated and accepted pilot into daily workflows, and when it prepares the workforce for the change. Phase 6, monitoring and governance, is when the firm measures returns over time and applies proportionate oversight. The phases are sequential, each one dependent on the output of the previous, and gated at every transition.

The phases are gated to address a specific deficit the literature locates between identifying an AI project and implementing it, the passage at which practical and academic knowledge thins [134]. Attaching a condition to each transition turns that gap into a sequence of explicit checkpoints, so that a firm never finds itself a step ahead of its own evidence. The route also assumes movement along a maturity gradient rather than a single leap: a G7 study sorts SME adopters into four ascending levels, namely AI Novices, Explorers, Optimisers, and Champions [36], and a firm entering as a novice is expected to complete one contained project and advance a single rung rather than emerge a champion in one pass. Given the financial and skills constraints the evidence ties to SMEs [135], the playbook deliberately adopts this single-step progression rather than assume a firm can sustain a transformative programme in one leap.

Figure 7.1 sets out the whole sequence at a glance: the six phases in order, the go or no-go gate at each transition, and the return to a fresh pass once a use case has been carried through.

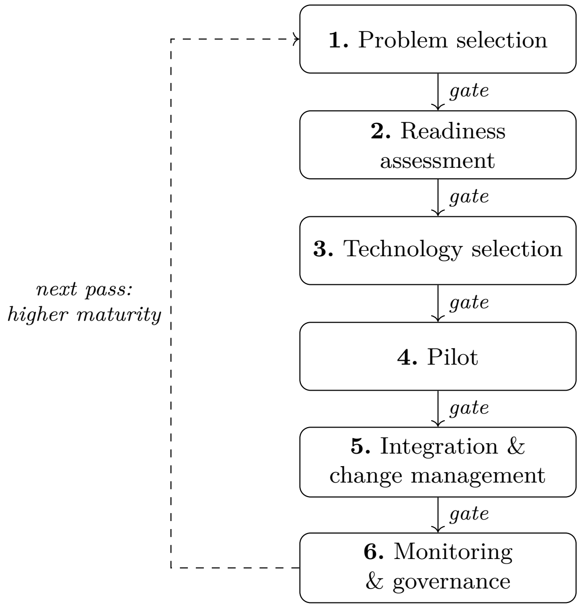


Figure 7.1. The six-phase gated adoption playbook. Each transition is an explicit go or no-go gate; a completed pass raises the firm's maturity and feeds the next.

7.2 Why a Playbook, and How It Was Built

The outline given below is called a playbook, not a method or a framework, and this decision is intentional. A framework or model is descriptive: it depicts what is true of a phenomenon. A playbook is executable and situational: it outlines the plays a practitioner performs, and which plays to deploy under which circumstances. The institutional source on which this chapter relies makes exactly this distinction: it calls the sequential, customisable guide the SME follows a playbook, and reserves the word framework for the scoring model embedded inside it [80]. This term fits the thesis audience as well: an owner-manager with few resources in time, money, and specialised expertise [135] who requires not an analytic map but an ordered series of plays and a rule for the conditions that justify each. It fits the structure too, since any single linear procedure would be a method, whereas the path varies by problem type, a different phase becoming the binding constraint in each (Section 7.4); a guide that prescribes distinct plays for different circumstances is, in short, a playbook.

It is a synthesis, and the two layers of the playbook make different claims to originality. The content of each phase is drawn from the literature, and every phase is supported by the citations where it is defined. The structure, however, is the contribution of this work:

the particular ordering of six gated phases; the precondition for each transition; and the assignment of a hardest phase to each problem category. None of the sources in the evidence base presents this pathway; it is built from different literatures, one focused on use-case selection, another on organisational readiness, others on procurement, piloting, change management, and governance. None spans the whole.

The development was carried out in a prescribed order, taking its cue from the gap the previous chapters identify, where failure is most concentrated: the passage from the decision to act on a specific project to its implementation [134]. This passage in particular was broken down into the smallest possible set of firm-level decisions and their sequence: what needs to be done, whether the firm is able to do it, whether it should be built or bought, whether it works, how it can be integrated, and how it can be kept safe. The corresponding SME-specific evidence, which supports each of those decisions as a discrete step and indicates how the firm should take it, was attached to its respective decision. Each transition was then checked against the evidence-based failures associated with that step, so that a checkpoint stands where firms are shown to stall. The final structure was checked for comprehensiveness against the five problem categories of Chapter 5 (the mapping is shown in Section 7.4).

Although original, the sequence is not arbitrary, since it follows process logic already recognised outside the SME setting. As an adoption rather than a construction process, its broad structure resembles the innovation-decision process of diffusion theory, which runs from first knowledge of an innovation, to the decision to adopt, to implementation, and on to confirmation [136, 137]. Its go or no-go transitions embody the gating principle of Stage-Gate, in which separate stages are bounded by gates that act as quality and go/kill checkpoints [138]. Its second phase draws on the technology, organisation, and environment framing through which firm-level adoption is conventionally assessed [139, 140]. What these general accounts leave under-specified is where the contribution for SMEs sits, since they describe adoption in the abstract. The playbook makes it concrete in three ways. It attaches to each transition a condition drawn from the specific evidence on where SME projects fail. It turns readiness from an assumption into an explicit gate. And it raises change management and governance, which such accounts fold into a single implementation step, to phases of their own. Table 7.1 records, for each phase, the evidence that motivates its inclusion.

Table 7.1. Each phase of the playbook and the evidence that motivates its inclusion as a distinct step.

Phase	Decision	Why it is a distinct step (motivating evidence)
1	Problem selection	A distinct, under-supported stage sits between identifying and implementing an AI project [134]; a first use case should be operational and measurable [29] and contained to internal support work [141], matched to a novice firm's off-the-shelf reach [36].
2	Readiness assessment	Organisational readiness, not the technology, decides the outcome [20]; firm-level adoption is located in technology, organisation, and environment [139, 140], with a concrete SME instrument available [142].
3	Technology selection	Identifying a suitable technology is itself the resource-constrained difficulty for SMEs [135], which tend to buy rather than build [143]; the choice should rest on total cost of ownership [144].
4	Pilot	Pilots are easy to build yet hard to scale [145], few pass the proof of concept to value [146], and a large share are abandoned for reasons a bounded pilot can pre-test [147].
5	Integration and change management	Most SMEs neither train staff nor set guidelines [65]; the workforce faces task transformation [148], and a person should stay in the loop on consequential output [149].
6	Monitoring and governance	Returns follow a multi-year productivity J-curve [63]; a scalable risk function suits firms of any size [150], and risk-tiered obligations apply to higher-risk uses [21].

7.3 The Six-Phase Adoption Playbook

7.3.1 Phase One: Selecting the First Problem

The first choice concerns which problem to take on, and the evidence points to operations and support rather than customer-facing or strategic functions [29, 141]. Measurability is the first criterion because value capture, not technical feasibility, is where AI efforts stall: roughly three-quarters of companies have yet to show tangible value from their AI, and what separates the rest is an explicit return-on-investment focus on every use case [145, 146]. A systematic review accordingly recommends that manufacturing SMEs begin with operational activities, on the grounds that operational outcomes are the easiest to measure through key figures [29], and an operational process yields such a figure almost by definition, whether documents processed, cost per transaction, or cycle time, each comparable before and after so the firm judges the result on evidence rather than impression. The logic holds beyond manufacturing, since any first use case should produce observable, quantifiable outcomes. A customer-facing or strategic use case, by contrast, makes a weaker first bet: its effect is harder to attribute cleanly, and in customer service it both fails more often, with users reporting no benefit nearly four times as often as from AI elsewhere [151], and carries documented paradoxes, such as customers left satisfied yet frustrated, that better models do not dissolve [152].

A second line of argument points to internal support processes as the first place to use AI. Guidance on this matter proposes business support processes as a smart sandbox for AI to be used to get early wins quickly and easily, before it is deployed in more client-facing functions [141]. Two attributes make them suitable for a first attempt: they are internal, so an early mistake stays inside the firm rather than reaching a customer, and they are repetitive, so the benefit from automation compounds quickly. Chapter 6 looks at one such example, document and process automation: large language models could pull out structured information from documents, invoices, and similar sources, while robotic process automation could handle the mechanical steps.

A narrow first scope is reinforced by how the least mature adopters behave. These adopters, which the OECD labels AI Novices, tend to use AI that is ready to run: standalone tools bought off the shelf, or AI features already built into software they own, rather than systems they assemble themselves from several AI components [36]. The first problem should therefore match the firm's level of experience rather than the apparent appeal of the use case: a novice that fixes on an ambitious decision-support use case, which needs data pipelines and predictive models, picks a problem no ready-made tool can yet solve for it. The disciplined move is to defer such a use case until later passes have raised the firm's maturity, and to begin with one whose solution is already packaged, so that capability accumulates safely instead of overreaching on the first attempt.

Selection is best treated as an explicit gate rather than an informal preference, since the literature locates a distinct, under-supported stage between identifying an AI project and implementing it [134]. Two criteria govern the choice. The first is measurability, the existence of a key figure such as processing cost per document or handling time per request against which the result can be judged [29]; the second is containment, a preference for internal support functions whose early gains stay inside the firm [141]. The right starting place is a use case that meets both and that a novice-level firm can serve with a single off-the-shelf product [36]; a firm selecting on any other basis, such as the visibility of a customer-facing application, takes on measurement difficulty and exposure beyond its current maturity.

7.3.2 Phase Two: Assessing Readiness and Maturity

Before adopting the selected tool the firm should assess whether it is prepared to absorb it, because organisational readiness, not the technology itself, frequently determines whether an SME deployment succeeds: the evidence ties the outcome of AI adoption to that readiness, which spans technological infrastructure, managerial expertise, and digital literacy [20]. The inclusion of managerial expertise alongside infrastructure is significant for the SME case, because the binding constraint identified throughout Chapter 5 is rarely hardware and frequently the absence of someone who understands both the business process and the technology well enough to direct the project.

Digital literacy is best read as a property of the workforce rather than of one individual. An estimate that six in ten European workers face task transformation from AI [148] suggests that the relevant literacy should be spread across staff rather than concentrated in a single technically minded employee. A readiness that rests on one champion is, by the same reasoning, fragile, since that person's departure would remove the capability; the assessment should therefore ask whether enough of the staff who will operate the tool possess the basic literacy to sustain it beyond any one person.

A practical, SME-specific instrument already exists: the European Commission's Digital Maturity Assessment [142], a free self-assessment run through the European Digital Innovation Hubs. It scores six dimensions, Digital Business Strategy, Digital Readiness, Human-Centric Digitalisation, Data Management and Connectedness, Automation and Artificial Intelligence, and Green Digitalisation, of which two carry the readiness verdict for a first use case: Human-Centric Digitalisation, which asks among other things whether the firm runs a staff-skills assessment, and Data Management and Connectedness, which asks whether it operates a data-management policy. These map onto the two constraints the literature flags, the digital-literacy gap and the data foundation. The data-management item carries particular weight: a generative tool can work on little proprietary data, but a predictive or retrieval system depends on the quality and organisation of the data it is given [90, 122], so a firm scoring poorly there should confine its first use case to one that tolerates thin data and treat building a data foundation as prerequisite work.

A more formal academic model sharpens this into a decision rule. It sets out five readiness dimensions, Strategic, Organisational, Technological, Capability, and Ecosystem readiness, and singles out foundational digital capability as the prerequisite on which the others rest [153]. The lesson for the SME is that readiness is gated by that foundation rather than averaged across the five: a firm weak on the basics is not ready however promising its strategy or its ecosystem.

A poor readiness result is best treated as prerequisite work, not as a reason to proceed regardless: because organisational readiness is what the evidence ties to success, adopting into an unready state raises the firm's own probability of failure [20], and the cost of closing the gap first, organising data or running a short literacy programme, is in most cases smaller than that of a failed deployment. Readiness is thus the second gate, forcing the firm to earn the right to adopt rather than assume it.

The output of the phase is a brief go/no-go verdict for that particular use case selected in Phase One, stating whether the company meets the basic data and skills requirements, identifying any gap that must be closed first, and noting the stage at which the firm stands on the maturity taxonomy. A firm that passes proceeds to technology selection; one that does not makes closing the identified gap its next step, because adopting into an unready organisation is precisely the pattern the readiness evidence associates with failure.

7.3.3 Phase Three: Selecting the Technology

With a use case chosen and readiness confirmed, the firm decides how to obtain the technology. For an SME the default is almost always to buy rather than build [143]. Identifying a suitable technology is itself a resource-constrained challenge for these firms: a synthesis of 106 papers concludes that wider AI use in SMEs depends on working through the difficulty of choosing suitable technologies, which their limited specialist resources make harder [135]. This argues for narrowing the field to packaged, ready-to-buy options that do not need the level of internal technical expertise that a customized buy or build option would demand.

The buy-default rests on the asymmetry in building capacity between large firms and SMEs. A review of 78 papers finds that large firms can build in-house AI while SMEs, viewing AI as expensive, fall back on simple low-cost solutions or outside consultants [143]; a survey corroborates the direction, reporting that 76 percent of deployed enterprise solutions were bought rather than built [94]. For an SME the realistic spectrum thus runs from consuming embedded AI to subscribing to an off-the-shelf product, with bespoke construction the exception [36]. The consultant route is legitimate where internal capability is genuinely absent, but it carries a total-cost caveat: a vendor-built solution can leave the firm locked in, dependent on that vendor for every later change and facing substantial switching costs [154], so the firm should retain ownership of its data and configuration.

The choice should rest on total cost, not headline price, following the total-cost-of-ownership logic that looks past the purchase price to the many other costs a purchase carries [144]. For an AI subscription these extend past the licence fee to four estimable items: the upfront integration with existing accounting, CRM, and storage systems; the data cleaning that makes the tool's input reliable; the staff training that Phase Five formalises; and the ongoing supervision that Phase Six installs. The discipline demands not precision but honest inclusion, because a build option that looks cheap on a salary basis can carry the ongoing maintenance burden that in-house machine-learning systems are known to accumulate [155], which a subscription externalises to the vendor. A documented case reports a 30 percent fall in maintenance costs from a predictive-maintenance deployment; this describes that case, not a generic pilot, and should be read as such [80].

For firms that cannot assess vendors unaided, a catalogue of vetted AI tools and the option to procure jointly with peers are described [80]: the catalogue pre-filters the field flagged as hard to navigate [135], while joint procurement, through sector associations or digital innovation hubs, restores some of the purchasing scale a lone small buyer lacks.

The hybrid should get a specific carve out, not a blanket approval. The hybrid (a combination of the subscription core and self-run component) assumes the capability that the resource-limited identification problem indicates most SMEs do not have (the ability to choose, implement, and sustain a reliable self-run piece) [135]. The hybrid is appropriate only when a specific need, such as a requirement that data not be sent to a third-party service, exists, and is not fulfilled by any of the packaged offers; and the verdict on readiness affirms this capability exists. Absent that need and capability, a hybrid only adds complexity and maintenance cost, so it remains a justified exception rather than the rule.

The output is a technology decision justified on total-cost grounds, consistent with the build-versus-buy default of Chapter 6 [94, 143]. In most SME cases this means subscribing to a packaged tool; a hybrid arrangement remains the exception that the Phase Two readiness verdict must explicitly support, not the base assumption.

7.3.4 Phase Four: Running the Pilot and Confronting Scale

The fourth phase is the bounded pilot, and the design should assume that the challenge is not the pilot itself, but rather how the pilot will become an at-scale solution. An overview of generative AI adoption reports that only 11 percent of organisations have taken it to scale, and observes that building a pilot is comparatively easy whereas turning it into an at-scale solution is a separate and harder problem [145]. The lesson for the SME, then, is that the bounded pilot should not be thought of as the solution itself, but rather as a constrained test case that specifically answers the question: can this use case transition to a production-like operation?

The scale of the gap shows in the evidence: only 26 percent of firms have moved past the proof of concept to generate value and only 4 percent lead in AI value creation [146], while at least 30 percent of generative AI projects are forecast to be abandoned after the proof of concept, blamed on poor data quality, inadequate risk controls, escalating costs, or unclear business value [147]. Those four causes double as a checklist of what a pilot must test for. That so few firms reach the leading edge is an argument for discipline, not a deterrent: the gains accrue to those who execute integration, measurement, and governance well rather than to those who run the most pilots, the leaders being those that put most of their effort into people and process rather than the algorithm itself [146], so an SME's differentiator is operational rigour rather than access to frontier models. Scaling, moreover, means something modest for a small firm, extending a validated capability from a test batch of invoices to all of them rather than an enterprise-wide rollout, and a use case chosen for its containment in Phase One is precisely one whose full scale the firm can reach.

A pilot earns its name by being bounded, and the boundary should be set on three axes. Its scope should cover one use case and one process rather than several at once, so that the result is attributable. Its duration should be long enough to expose the tool to representative variation, such as a full billing cycle for an invoice process, yet short enough that an unsuccessful trial is cheap to abandon. Its reversibility should be preserved, so that the firm can withdraw without disrupting the underlying operation if the trial fails. These constraints follow from the observation that the easy part is the pilot and the hard part is the transition beyond it: a pilot kept deliberately small and reversible is the safest way to learn whether that harder transition is worth attempting [145].

The remedy is to declare success criteria before the pilot begins, against the key figure chosen in Phase One [29]. Each recorded abandonment cause then becomes a concrete check: data quality, by running the pilot on live rather than curated data; risk control, through the human-oversight design of Phase Five; escalating cost, by extrapolating the pilot's measured per-unit cost; and unclear value, by the pre-declared threshold itself [147]. A pilot instrumented to answer these four questions yields a go or no-go decision grounded in the same factors the evidence ties to success or failure.

The pilot should also be read against how SMEs actually use AI: a D4SME survey finds that 61 percent of SMEs report using it, yet some 76 percent remain novices applying off-the-shelf tools to isolated tasks [129]. A first pilot that delivers one isolated capability well is therefore a success on its own terms; connecting it to surrounding systems is the work of Phase Five, not a burden on the pilot.

A tested capability and a documented go/no-go call are the outputs of this phase. Where the pilot has met its pre-specified requirements, the organization can move on to integration. Where it has not, the organization should stop or re-scope (evidence that the majority of companies stay at a pilot stage implies that an honest negative result is the common case rather than an anomaly [145, 146]). If the pilot is allowed to continue without the pre-specified requirements having been fulfilled, it is a candidate for the recorded post-pilot abandonment [147].

7.3.5 Phase Five: Integration and Change Management

A validated pilot delivers value only once it is woven into workflows and the workforce is ready to use it, and on both counts SMEs are under-invested: a third or fewer train staff or set internal guidelines, and 49.8 percent report that employees lack the right skills [65]. The technical side means taking the capability out of isolation, the shallow, isolated-task use recorded in the survey being exactly what integration must overcome [129]: an invoice-extraction model adds little if its output is re-keyed by hand, and value arrives only when it flows automatically into the downstream system. The Phase Three buy-default eases this: packaged tools commonly arrive with ready-made connectors to the systems a firm already runs, which is why technology selection counted integration effort as a cost in its own right.

Training is best seen as ongoing rather than a one-off, because the work itself is changing: six in ten European workers face task transformation from AI, which frames upskilling and reskilling as the response [148]. This gives the SME a constructive narrative, staff being equipped with a more capable tool rather than edged toward redundancy, and it argues for AI literacy spread across the affected personnel rather than concentrated in one champion, so the deployment survives that champion's absence or a tool upgrade. Since a small firm has little slack to release staff for instruction, the training should be short and specific to the deployed use case, widening only as further use cases are added.

Resistance must be managed, not ignored. Workforce resistance, including fear of job displacement, alongside limited financial and human resources and weak infrastructure, is identified as a barrier to AI in SMEs [156]: where staff fear replacement, that resistance can undermine a deployment, so pairing training with the Phase One framing of AI as an automator of routine internal steps lets the firm present roles as changing rather than disappearing [141]. The same resource constraint argues for incrementalism, integrating one workflow at a time and stabilising it before the next, since integrating broadly and training everyone at once would overwhelm a small firm and risk a stalled, half-adopted deployment.

Training and written guidelines are paired deliberately, because each covers a weakness of the other: training builds the skill to operate the tool, while guidelines define the bounds of acceptable use. The data show most SMEs do neither, and doing only one leaves either capable-but-ungoverned use or rules staff cannot follow [65]. For a firm without a compliance function, a short written statement of permitted and prohibited uses is a low-cost control that also supplies the documentation the governance phase will require.

The integration design should keep a person in the loop, especially on any consequential output. A human-in-the-loop design keeps human control, feedback, and judgement present at points along the AI pipeline rather than leaving the model to act alone [149]. For an SME, for example, using a generative language model to draft a document or a predictive model to flag exceptions, the design is simple: the model proposes and a person disposes, reviewing and deciding before anything acts. One design goal is to prevent errant output reaching a customer; the second is to allay displacement anxieties, so employees are not redundant, but in an oversight role [156]. In addition, this human oversight is exactly what the later governance phase will require for higher risk uses, so the change management decision and the subsequent requirement of compliance is the same action.

Integration and training are also the complementary investment on which the return depends: the productivity literature on general-purpose technologies holds that gains arrive only after a firm makes the surrounding investments in process change and skills [63], so their cost falls in Phase Five before the benefit becomes visible in Phase Six. The output of the phase is therefore a deployment connected to its workflows, a trained workforce, written usage guidelines, and a defined point of human oversight, precisely the measures the data shows most SMEs omit [65, 129], which is what converts a successful pilot into an operating capability rather than a demonstration the firm cannot sustain.

7.3.6 Phase Six: Monitoring and Governance

The final phase installs measurement and oversight, beginning with a realistic expectation of when returns appear. The evidence describes a productivity J-curve for general-purpose technologies such as AI: because they require complementary investments that are intangible and poorly measured, such as the process change and skills of Phase Five, measured productivity dips before it rises [63]. For the SME this means return on investment must be read over several years rather than a single quarter, since an early measure taken during the investment phase understates the eventual return and can prompt a premature shutdown of a deployment that was on course to repay.

Measurement is best framed as a continuing function rather than a one-off audit, and the NIST AI Risk Management Framework supplies a fitting structure. One of its four functions, Measure, analyses, benchmarks, and monitors AI risk, and the framework is explicitly voluntary and for organisations of all sizes, which makes it legitimately applicable to an SME [150]. For the firm it reduces to mixing a slow measure with fast ones: return on investment read over the multi-year, J-curve horizon and not over-interpreted in any quarter, alongside the faster health signals the abandonment evidence flags, drift in input-data quality, creep in per-unit cost, and the rate at which human reviewers override the tool [147]. A climbing override rate is an early warning that reliability has dipped, the kind of live signal that machine-learning systems require continuous monitoring to catch [155], letting the firm intervene before the slow value measure turns negative.

Governance should be proportionate, and the risk tiers of the EU AI Act let an SME concentrate effort where the stakes are highest: the Act sorts uses into four levels, unacceptable, high, limited, and minimal risk, and attaches to the high-risk tier obligations of risk assessment, logging, human oversight, and documentation [21]. The first governance task is therefore to classify the use case: most SME first use cases, such as drafting internal documents, fall outside the high-risk tier and need little beyond the Phase Five guidelines, while a high-risk use must meet the fuller obligations. Several of these are already in hand from earlier phases, the human oversight from the integration design, the documentation from the usage guidelines, and the logging from the monitoring routine, so reusing those artefacts keeps the compliance burden proportionate for a firm with no dedicated legal function.

A monitored, integrated, and governed first deployment lifts a firm out of the shallow pattern the D4SME survey records, where most adopters stay beginners applying off-the-shelf tools to isolated tasks [129]. A firm that has carried one use case through all six phases has broken that ceiling for it, and the experience gained in selecting, assessing, piloting, integrating, and governing is itself a complementary asset that lowers the cost of the next pass. Each completed deployment thus delivers its own value and raises the firm's readiness for a more demanding use case.

The phase ends with a continuous, light-touch routine: the key metric read against the J-curve, the Measure function applied proportionately, and governance intensity set by the use case's risk tier [21, 63, 150]. With this in place the firm has completed one full pass of the playbook.

7.4 Applying the Playbook by Problem Category

The six phases apply to every category, but they do not weigh the same across them: for each problem category a different phase becomes the bottleneck, the one that decides whether the project succeeds. What follows names that hardest phase for each of the five classes of Chapter 5 and sets out the evidence behind it, drawing on the solutions already mapped in Chapter 6 rather than re-deriving them; Table 7.2 summarises the result. The hardest phases then settle into an order of adoption.

Document and process automation. This is the one category with no genuinely hard phase. Its output is internal and countable and its tool market is mature, so even the step that asks for most judgement, the first, choosing a process with a measurable figure, stays light, and a packaged product with plug-and-play connectors handles the rest. The maturity shows in the numbers: the intelligent document processing market is set to grow from \$2.30 billion in 2024 to \$12.35 billion by 2030, fastest in its SME segment [98]; an invoice costs \$9.40 to process against \$2.78 for best-in-class firms, takes 9.2 days, and clears untouched only 32.6 percent of the time [99]; and even among smaller firms adoption here is comparatively strong. For these reasons it is the standard place to begin.

Customer service. The start is easy but the finish is not. The data barrier is low, since 37.3 percent of the SMEs that use generative AI already apply it here [65] and packaged tools provide the around-the-clock coverage once reserved for large firms [106], yet the hardest phase moves to the fifth, integration and human oversight, because what the system produces reaches the customer directly. The exposure is not hypothetical: AI in customer service fails at about four times the rate of AI used elsewhere, and nearly one in five users report no benefit at all [151], while peer-reviewed work catalogues six paradoxes of generative customer service that better models do not dissolve, among them higher quality yet less empathy, and greater power yet greater vulnerability [152]. The effort therefore concentrates on the escalation path that routes uncertain cases to a person [29], not on choosing the tool.

Marketing and sales. This category is two tasks in one, with opposite barriers, so the first job is to tell them apart. It is also the function where SMEs adopt AI most: across the EU, 34.7 percent of AI-using firms apply it to marketing or sales [77], and 49.7 percent of the firms using generative AI direct it at marketing content [65], precisely because one of its two tasks is so easy to begin. That task is generative, the production of copy, posts, and campaign material, and it needs almost no proprietary data; it usually arrives bundled into a marketing platform the firm already runs [112], so the build-versus-buy question resolves by default to using what is already paid for. The return is measured, a randomised trial recording a 40 percent reduction in time and an 18 percent gain in quality on a writing task [16], and it falls most generously to the smallest firms, which see some of the strongest returns on blog content [116]. The second task, predictive work such as lead scoring, is the mirror image: it cannot run without a history of customer records, so its bottleneck is the same data-readiness phase that constrains decision support. A modest dataset can already perform, one study classifying leads correctly 98.39 percent of the time across 23,154 records [113], but the harder obstacle is access rather than volume: a firm that runs its marketing through a platform frequently does not own the customer data the platform holds, so the very records the predictive model needs may never reach its hands [19]. The category therefore has no single hardest phase; what it demands first is the judgement to recognise which of the two tasks a use case really is, since the generative one can be bought off the shelf while the predictive one waits on data the firm may not control.

Knowledge management. The bottleneck is the second phase, readiness, since retrieval over a firm's own documents is only as good as the corpus beneath it. Most SMEs are not ready: their documents sit scattered across email, file shares, and messaging logs [85], and the infrastructure, skills, and budget that knowledge work demands are exactly what the literature finds wanting in smaller firms [122]. Retrieval-augmented generation, which feeds a generative model passages pulled from the firm's own documents, is still immature: most deployments lean on default building blocks rather than hardened custom ones, with 63.6 percent resting on GPT models and 80.5 percent on out-of-the-box pipelines [84], and a corpus left unmaintained degrades the answers without ever signalling an error [90]. The first task is to organise and keep the data, and only then to choose a tool.

Decision support. The bottleneck is again readiness, and here it bites hardest. The adoption gap is at its widest, 17.4 percent of small firms against 52 percent of large ones [36], because predictive models depend on organised historical data that the least-mature firms have not assembled. The barrier is not absolute, a churn model has reached 91.25 percent accuracy on just 4,911 records [128] and the reward is documented, AI-adopting SMEs being 15.1 percentage points more likely to report revenue growth above 30 percent [66], but the signal is slow, a forecast proving out only over weeks, which makes this a second or third move rather than a first.

Table 7.2. The hardest phase of the playbook for each problem category, and why.

Category	**Hardest phase**	**Why**
Document and process automation	None acute (Phase One, selection)	Output countable and internal; mature off-the-shelf market
Customer service	Phase Five, integration and oversight	Output reaches customers; high customer-facing failure rate
Marketing and sales	Varies by task	Generative needs no hard phase; predictive shares the data barrier
Knowledge management	Phase Two, readiness	Retrieval needs an organised document corpus
Decision support	Phase Two, readiness	Predictive models need historical data; slow signal

These hardest phases settle into an order. Document automation and customer service can be taken on first, because their bottlenecks, a light first selection and a manageable oversight design, are within reach of today's packaged tools; knowledge management and decision support are better deferred, because they share a bottleneck, a data foundation the firm does not yet possess, and that is precisely where low-maturity adopters tend to stall [129]. The sequence orders the work of Chapter 6 without altering it.

7.5 The Playbook in Practice: Automating Accounts Payable

To illustrate the playbook in action rather than in the abstract, this section applies the six phases to one use case: the automation of accounts payable, that is, the processing of supplier invoices, which the category analysis of Section 7.4 identifies as the natural place to begin. The exercise is a constructed application rather than a field report, and it is parameterised throughout with published figures rather than assumed ones. The firm is a representative small manufacturer of around forty employees, at the lower edge of the size range where adoption lags most: across the OECD only about a fifth of firms with 50 to 249 employees, and a smaller share of the smallest, use AI, against two-fifths of large firms, and the same report records a real micro-wholesaler in Japan that already runs AI agents across inquiries, invoicing, and shipping [36]. The use case itself is likewise proven in practice: in a documented corporate deployment, the utility group Veolia paired an AI document-extraction engine with software robots for invoice processing and reported an 87 percent gain in processing efficiency [141].

Problem selection. Phase One admits a use case only if it is measurable and contained (Section 7.3.1), and invoice processing meets both criteria comfortably. It carries a clean key metric: an average accounts-payable function reports a cost near $9.40 to process one invoice over a cycle of about 9.2 days, with under a third of invoices clearing without human intervention [99], so the outcome can be judged on data rather than impression. It is likewise internal and repetitive, so an early error remains inside the firm while the gains from automation accumulate, and it is served by a packaged market within reach of a novice firm, intelligent document processing being among the fastest-growing software segments for smaller firms [98].

Readiness assessment. Phase Two grants passage only if the data and the staff can support the tool (Section 7.3.2). For this use case that is less a question of hardware than of whether the invoice data is organised and the accounts staff hold the literacy to oversee an automated flow, and the digital-maturity items on data management and staff skills reveal any gap [142]. Document automation tolerates a thin start, since its inputs are the firm's own structured invoices rather than a large proprietary dataset, so a firm that scores adequately may proceed, while one that does not schedules that work first.

Technology selection. Phase Three defaults to buying and prices the decision on total cost of ownership (Section 7.3.3). The firm therefore subscribes to a packaged tool that integrates document extraction with workflow automation, the default the evidence supports for a resource-constrained SME [143], in a purchase-to-pay market mature enough that a documented provider deploys its software robots within two to four weeks [157]. The full costing looks past the licence fee [144] to the integration with the existing accounting system, the data cleaning, the training required in the next phase, and the ongoing oversight required in the last, because a build option that looks cheap on a salary basis would carry the maintenance burden that a subscription externalises [155].

Pilot. Phase Four demands success criteria declared before the trial begins, on a scope that is bounded and reversible (Section 7.3.4). The firm runs the tool for a single billing cycle of live invoices and fixes the criterion in advance against the phase-one figure: cost and cycle time must fall toward the best-in-class marks of about $2.78 an invoice and three days that the same benchmark records [99]. Running on live invoices rather than curated ones tests each of the four documented causes of post-pilot abandonment, namely poor data quality, weak risk controls, escalating cost, and unclear value [147], before any commitment to scale. That a disciplined back-office deployment can repay the effort is itself documented, if at enterprise scale: the software robots of Telefónica O2 reached payback within twelve months and a three-year return between 650 and 800 percent [158].

Integration and change management. Phase Five prescribes connection to the downstream workflow, training paired with written guidelines, and a defined point of human oversight (Section 7.3.5). A pilot that meets its threshold is therefore woven into the accounting workflow so that extracted data flows on without re-keying, and the accounts staff are trained and given a short written statement of permitted use, the paired controls most SMEs are shown to omit [65]. A person stays in the loop on the exceptions the tool cannot clear [149]; at the benchmark such exceptions, those flagged for missing data or coding errors rather than merely touched for routine handling, run at roughly one invoice in seven [99]. That point of oversight both keeps an errant payment from leaving the firm and recasts the affected role as supervision rather than redundancy, the framing the resistance evidence recommends [156].

Monitoring and governance. Phase Six sets the measurement horizon and the governance tier (Section 7.3.6). The firm reads its return over a multi-year horizon, since the benefits of a general-purpose technology trail the investment that produces them [63], while tracking faster signals such as drift in extraction quality, creep in cost per invoice, and the rate at which reviewers override the tool. Governance stays proportionate: invoice automation falls outside the high-risk tier of the EU AI Act, so little is required beyond the phase-five guidelines and the oversight already in place [21].

One pass therefore carries a single, measurable, internal use case from selection through to a governed capability, and each pass lowers the cost of the next. The pass shown here is constructed, but the deployments cited alongside it are not, and together they indicate that the route is practicable rather than merely plausible. The same discipline also signals when not to proceed: a firm that instead pursued demand forecasting would meet a different verdict at the readiness gate, since predictive use cases depend on organised historical data, where the small-to-large gap is widest (Section 7.4). The binding constraint is not data volume, as a predictive model performs on only a few thousand records [128], but data the firm has not yet organised or does not control, so the disciplined response is to defer the use case and build the data foundation first, the gate doing its work.

7.6 The Agentic Shift and Future Evolution

The playbook draws on AI today that an SME can buy, yet is headed toward agentic systems that do not only respond to a prompt, but also can invoke tools, make decisions, and complete a task autonomously over a series of steps [13]. This is framed as a transition to adaptive, multi-stage, end-to-end execution [72], the next stop on the arc from automation based on rules to cognitive enhancement, covered in Chapter 2. The question for SMEs is not whether this trend is coming, but whether it alters the systematic approach outlined in the playbook.

The projected scale of this shift is large. Task-specific agents are expected in 40 percent of enterprise applications by 2026, up from under 5 percent in 2025 [131]; by 2029, agentic AI is projected to resolve 80 percent of routine customer-service issues on its own, bringing customer-service costs down by 30 percent [109], and the market opportunity for vertical, industry-specific agents is valued at one to two trillion dollars [54]. The present capability is far smaller than the forecast. On TheAgentCompany benchmark, which sets agents real consequential office tasks, even today's best agent model only autonomously solves around 30 percent of the tasks [159], and most organisations that have invested in AI have yet to translate it into tangible value [146]. The direction of travel is set, but the capability still lags well behind the projections.

None of this changes the playbook for an SME, because an agentic tool removes none of the barriers the six phases were built to address; those barriers belong to the firm, not to the model it buys. The base rates leave little room for optimism. More than 80 percent of AI projects fail, twice the rate seen outside AI [160], and when SMEs explain why they have not adopted, the reason named most often is a plain lack of expertise, given by 70.9 percent of non-adopters [28]. That shortfall is structural rather than incidental, since over half of the people working in SMEs do not hold sufficient digital competence [59]. The pattern of under-investment compounds it: roughly half of the firms that have adopted AI add nothing in training, talent, or software, and on robotic process automation, the rule-based ancestor of agentic action, they already lag larger firms by almost 17 percentage points [38]. A tool that asks for more autonomy therefore lands in a wider readiness gap, not a narrower one, and aimed at a firm that is not ready it is more likely to fail, not less.

Such caution is the prudent position. Agents still make too many errors to be trusted with business processes of real financial value, and reliable autonomy remains several years out, conditional on progress in error rates, cybersecurity, and keeping system behaviour within intended bounds [13]. The better strategy is not to wait but to keep the option alive: run a controlled use case through the complete six-phase arc with today's stack, and preserve the technical openness that avoids lock-in, favouring exposed APIs, event-driven integration, and exportable structured data [72]. The design choices behind that stance are set out in Chapter 6 (Section 6.7.2); what the playbook actually adds is the sequencing principle, that agentic capability is the second pass, after the six-phase muscle has been built and there is a need for multi-step autonomy in a specific use case.

7.7 Implementation Checklist

Ultimately, the playbook reduces to a small set of go or no-go questions that an owner-manager can run through at each gate, deciding whether the firm is ready to proceed or should pause and close a gap first. The checklist below synthesises the six phases into that decision form, drawing only on the reasoning already established in the preceding sections.

- Phase One, problem selection. Does the chosen use case have a key figure, such as cost per document or handling time, against which a result can be judged? Is the process internal enough that an early failure stays inside the firm? Can a single off-the-shelf tool serve it without custom integration?

- Phase Two, readiness assessment. Do the staff who will operate the tool possess the basic literacy to use it, beyond a single champion? Is the data the use case depends on organised well enough to support it? If either answer is no, is the gap scheduled as prerequisite work before adoption?

- Phase Three, technology selection. Has the option been costed on a total-cost basis, including integration, data preparation, training, and ongoing oversight, rather than on licence price alone? Is the default packaged purchase being chosen unless a specific constraint justifies building or a hybrid?

- Phase Four, pilot. Are success criteria declared in advance against the Phase One key figure? Is the pilot bounded in scope, duration, and reversibility, and run on live rather than curated data? Does the result meet the threshold to proceed, or is the disciplined response to stop or re-scope?

- Phase Five, integration and change management. Is the validated capability connected to the systems where work happens, rather than left as an isolated task? Have the affected staff been trained and written usage guidelines issued? Is a point of human oversight defined for any consequential output?

- Phase Six, monitoring and governance. Is the return measured over a multi-year horizon rather than judged in one quarter? Are health signals, such as data drift, unit cost, and the human override rate, monitored continuously? Is the governance intensity matched to the use case's risk tier rather than applied uniformly?

If a company can answer yes to the questions in each phase, or if it knows what work needs to be done to answer yes, and has it on the calendar, the company is working through the phases in the order the evidence suggests. A no at any gate is not a stop signal but an instruction to close the gap first, which is the discipline the whole playbook enforces.

7.8 The Playbook in Summary

The six phases form a single route from an unaddressed problem to a governed, operating capability. Table 7.3 consolidates it, recording for each phase the action the firm performs and the documented finding that anchors it; the table reads top to bottom as the firm would move, and while each entry rests on the sourced evidence developed in the corresponding section, the gated sequence that binds them into a single route is the synthesis set out in Section 7.2.

The path is purposefully cautious, because caution is the point. Each stage is a hurdle, and the evidence shows that SME attempts most often fail at the transitions between stages, which is exactly where the playbook places a gate [134, 147]. Take a firm that steps through the barriers sequentially. It chooses a specific problem that is measurable and has boundaries, it validates its readiness, it commits based on life-cycle cost, it tests against the pre-specified success metrics, it puts the tooling in place with associated training and controls, and it tracks its performance on a sensible timescale. A firm that takes such a path has done what the evidence supports at each step for a company like it. After one full pass the firm is more mature than it began, so the playbook describes not a single project but a repeatable method for adopting AI one contained problem at a time.

These claims are bounded by several limits, and naming them is part of using the playbook honestly. It is an evidence-based construct rather than a tested product: it has not yet been validated on a live SME, and the evidence behind its earliest phase is drawn largely from manufacturing and extrapolated with care to other settings [29]. Its scope, moreover, is the firm that adopts AI through the as-a-service market rather than one building its own models, and the move toward agentic systems set out in Section 7.6 will require the sequencing rule to be re-examined as that capability matures. The natural

Table 7.3. The six-phase SME adoption playbook, with the documented finding anchoring each phase.

Phase	Action	Anchoring evidence
1	Problem selection	Manufacturing SMEs are advised to start with operational activities, which are easiest to measure by key figures [29]; support processes serve as a quick-win testing ground [141]; AI novices rely on embedded or off-the-shelf solutions [36].
2	Readiness assessment	Organisational readiness, including managerial expertise and digital literacy, is crucial to adoption success [20]; the EC Digital Maturity Assessment scores six dimensions including staff skills and data management [142].
3	Technology selection	Large firms build while SMEs buy or rely on consultants [143]; total cost of ownership looks beyond purchase price [144]; 76 percent of deployed enterprise solutions were bought [94].
4	Pilot and scaling	Only 11 percent of firms reach scale with generative AI [145]; only 26 percent pass the proof of concept to value [146]; at least 30 percent of projects are abandoned after the proof of concept [147].
5	Integration and change management	A third or fewer of SMEs train staff or set guidelines, and 49.8 percent report a skills gap [65]; six in ten workers face task transformation [148]; human-in-the-loop keeps oversight in the pipeline [149].
6	Monitoring and governance	The productivity J-curve means returns lag and must be measured over years [63]; the NIST MEASURE function monitors risk for organisations of all sizes [150]; the EU AI Act sets four risk tiers [21].

next step is therefore to carry a single use case through all six phases on an operating firm and to measure whether the gated structure delivers what it claims.

Chapter 8

Conclusion

This thesis originated in an existing operational void: while AI-as-a-service has become viable for small and medium-sized enterprises, there remains a distinct lack of guidance on the problems that require prioritising, what technologies should be applied to each problem, and how to approach an implementation path that allows a company to operate AI services sustainably. The foregoing chapters fill this void by charting an arc from the historical barriers to SME AI adoption through to a breakdown of possible problem-types to which AI has a solution and an accompanying breakdown of possible solutions, and finally through to a sequence of steps to deploy an existing service as AI-powered, one problem at a time, with governance over the process of doing so.

The first research question concerns which categories an SME should prioritise, and what makes one easier to enter than another. The barriers to adoption have dropped steeply and unevenly (Chapter 4): the cost, infrastructure, and integration hurdles that once kept smaller firms out have dissipated, with the cost of inference in particular dropping roughly a thousandfold in three years, while the barriers that remain, organisational rather than technical, persist, with a lack of expertise the reason most often cited for not adopting. In this setting the problem space resolves into five evidence-based categories, arranged by an ascending entry barrier (Chapter 5): document and process automation stands lowest, its results internal and countable and its software well established, while decision support stands highest, gated by a data foundation that the widest small-to-large adoption gap shows most firms lack. What sets the height is how far a category depends on proprietary data, how exposed its outputs are, and how tightly its results must be integrated; and beneath all five lies one common constraint, the firm's own skills and knowledge.

The second question examined what solution patterns and tool families addressed those categories, and which an SME would be best advised to buy, bundle, or build. Buy or bundle overwhelmingly is the default (Chapter 6): the proportion of enterprise AI solutions bought rather than built is now approximately three-quarters, and probably higher in SMEs. Each category aligns to a well-known solution pattern on a four-rung complexity ladder where the first two, embedded tools and packaged tools, match the majority of SMEs; only unusual firms with unusual data or regulations, or unusual technical requirements, should build. One should choose to build, bundle, or buy on the basis of cost of ownership, not list price, since a build that looks cheap when calculated in terms of one annual salary can carry ongoing maintenance and reliability costs that a subscription externalises to the vendor.

The third research question asks how an SME should structure its staged adoption approach so that sustainability is achieved. The answer is a six-phase, gated playbook (Chapter 7): problem selection, readiness assessment, technology selection, bounded pilot, integration and change management, monitoring and governance; at the end of each phase, an explicit go or no-go decision is made. The design is intentionally conservative; the evidence indicates that SME failures are not due to the technology but at the boundaries between problem selection and implementation: only a small proportion of SMEs succeed in advancing pilot results to value and a considerable fraction of generative AI projects are terminated after proof-of-concept completion. Explicitly naming each boundary, and defining conditions for its completion, means that this gap is no longer unmanaged; it becomes a series of checkpoints for a firm to pass through, such that it is never a step ahead of the evidence for a given stage. The playbook is a repeatable method rather than a one-off: every pass through the phases advances maturity and reduces the costs of the next iteration.

The thesis makes two distinct contributions. On a theoretical level, it categorises SME AI challenges, drawing on triangulated findings, by the scale of the barrier each imposes and conceptualising adoption obstacles as reduced in severity but altered in form. On a practical level, it synthesises this analysis into a decision-making framework that a non-AI-specialist owner-manager can use to follow through a clear progression of steps; in doing so, it consolidates a diverse body of research into a coherent process.

These conclusions are subject to several constraints. The bulk of data underpinning the earliest phase is focused on manufacturing and it has been carefully extrapolated from the available data. Many of the market and value metrics defining the opportunity are sourced from vendor and advisory reports rather than academic sources and these sources are noted for that purpose. Crucially, the taxonomy and playbook are not tested products but are rather evidence-based constructs and have not been empirically evaluated on a living SME; the taxonomy is also scoped for now and will evolve as agentic capabilities mature.

This maturation is the principal direction for future work. Agentic AI, treated here as a deliberate second pass, is advancing rapidly in capability but lags in reliability, with only about one-third of consequential real-world office tasks completed by the most capable agents at present; as that reliability improves, the playbook's sequencing rule will itself need re-evaluation. The most valuable next step is to validate the playbook on an actual SME, taking a single use case through all six phases and measuring whether the gated structure delivers what it claims, so as to convert AI from a demonstration a firm cannot sustain into an operating capability it can.

The larger significance of this is that the technology is no longer the barrier to AI in the SME. The tools are there, and for most SME problems they can be bought; the scarce thing is a disciplined staged approach to adopting them. It is this process, not any tool, that the thesis has attempted to deliver.

Bibliography

[1] A. M. Turing. «Computing Machinery and Intelligence». In: *The Essential Turing.* Ed. by B. J. Copeland. Originally published in Mind, Vol. 59, No. 236, pp. 433–460 (1950). Clarendon Press, Oxford, 2004. DOI: 10.1093/oso/9780198250791.003.0017.

[2] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon. «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence». In: *AI Magazine* 27.4 (2006). Reprint of 1955 proposal.

[3] R. Cordeschi. «AI's Half Century: On the Thresholds of the Dartmouth Conference». In: *IA Retrospettiva, AI*IA (Italian Association for AI)* III.1/2 (2006), pp. 109–114.

[4] J. M. Mira. «Symbols Versus Connections: 50 Years of Artificial Intelligence». In: *Neurocomputing* 71.4-6 (2008), pp. 671–680.

[5] J. Durkin. «Expert Systems: A View of the Field». In: *IEEE Expert* 11.2 (1996), pp. 56–63.

[6] J. Lighthill. *Artificial Intelligence: A General Survey.* Tech. rep. Science Research Council, London, 1973.

[7] D. Michie. «Lighthill 17 Years On: End of a Shotgun Divorce». In: *The Knowledge Engineering Review* 5.4 (1990), pp. 277–293.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton. «ImageNet Classification with Deep Convolutional Neural Networks». In: *Advances in Neural Information Processing Systems (NeurIPS).* Vol. 25. 2012.

[9] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. «Attention Is All You Need». In: *Advances in Neural Information Processing Systems (NeurIPS).* Vol. 30. 2017.

[10] L. Vinsel. «Don't Get Distracted by the Hype Around Generative AI». In: *MIT Sloan Management Review* (2023). URL: https://sloanreview.mit.edu/article/dont-get-distracted-by-the-hype-around-generative-ai/.

[11] M. I. Jordan and T. M. Mitchell. «Machine Learning: Trends, Perspectives, and Prospects». In: *Science* 349.6245 (2015), pp. 255–260. DOI: 10.1126/science.aaa8415.

[12] Y. LeCun, Y. Bengio, and G. Hinton. «Deep Learning». In: *Nature* 521 (2015), pp. 436–444. DOI: 10.1038/nature14539.

[13] T. H. Davenport and R. Bean. «Five Trends in AI and Data Science for 2026». In: *MIT Sloan Management Review* (2026). URL: https://sloanreview.mit.edu/article/five-trends-in-ai-and-data-science-for-2026/.

[14] M. Chui, E. Hazan, R. Roberts, A. Singla, K. Smaje, A. Sukharevsky, L. Yee, and R. Zemmel. *The Economic Potential of Generative AI: The Next Productivity Frontier*. Tech. rep. McKinsey Global Institute, 2023. URL: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier.

[15] E. Brynjolfsson, D. Li, and L. R. Raymond. «Generative AI at Work». In: *The Quarterly Journal of Economics* 140.2 (2025), pp. 889–942. DOI: 10.1093/qje/qjae044.

[16] S. Noy and W. Zhang. «Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence». In: *Science* 381.6654 (2023), pp. 187–192. DOI: 10.1126/science.adh2586.

[17] D. Luitse. «Platform Power in AI: The Evolution of Cloud Infrastructures in the Political Economy of Artificial Intelligence». In: *Internet Policy Review* 13.2 (2024). DOI: 10.14763/2024.2.1768.

[18] S. Lins, K. D. Pandl, H. Teigeler, S. Thiebes, C. Bayer, and A. Sunyaev. «Artificial Intelligence as a Service». In: *Business & Information Systems Engineering* 63.4 (2021), pp. 441–456. DOI: 10.1007/s12599-021-00708-w.

[19] D. Cutolo and M. Kenney. «Platform-Dependent Entrepreneurs: Power Asymmetries, Risks, and Strategies in the Platform Economy». In: *Academy of Management Perspectives* 35.4 (2021), pp. 584–605. DOI: 10.5465/amp.2019.0103.

[20] Y. Yesuf and Z. Fields. «Artificial Intelligence Adoption as a Driver of Innovation and Competitiveness in SMEs: A Bibliometric and Systematic Review». In: *F1000Research* 14 (2025), p. 1187. DOI: 10.12688/f1000research.171494.1.

[21] European Commission. *Regulatory Framework on Artificial Intelligence*. European Commission, Shaping Europe's Digital Future. 2024. URL: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

[22] European Commission. *SME definition*. European Commission website. 2025. URL: https://single-market-economy.ec.europa.eu/smes/sme-fundamentals/sme-definition_en.

[23] T. Gibson and H. J. van der Vaart. *A Universal Definition of Small Enterprise: A Procrustean Bed for SMEs?* World Bank Group blog. 2008. URL: https://blogs.worldbank.org/en/psd/a-universal-definition-of-small-enterprise-a-procrustean-bed-for-smes.

[24] Eurostat. *Use of Artificial Intelligence in Enterprises*. Statistics Explained, European Commission, Luxembourg. 2025. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises.

[25] A. Madgavkar, M. Piccitto, O. White, M. J. Ramírez, J. Mischke, and K. Chockalingam. *A Microscope on Small Businesses: Spotting Opportunities to Boost Productivity.* Tech. rep. McKinsey Global Institute, McKinsey & Company, 2024. URL: https://www.mckinsey.com/mgi/our-research/a-microscope-on-small-businesses-spotting-opportunities-to-boost-productivity.

[26] S. Raes. *Understanding SME Heterogeneity: Towards Policy Relevant Typologies for SMEs and Entrepreneurship.* Tech. rep. OECD SME and Entrepreneurship Papers No. 28, OECD Publishing, Paris, 2021. URL: https://www.oecd-ilibrary.org/economics/understanding-sme-heterogeneity_c7074049-en.

[27] M. Manjon, Z. Aouni, and N. Crutzen. «Untangling the Relationship Between Small and Medium-Sized Enterprises and Growth: A Review of Extant Literature». In: *International Entrepreneurship and Management Journal* 20.2 (2024), pp. 727–760. DOI: 10.1007/s11365-023-00830-z.

[28] OECD. *The Digital Transformation of SMEs.* Tech. rep. OECD Studies on SMEs and Entrepreneurship, OECD Publishing, Paris, 2021. URL: https://www.oecd.org/en/publications/the-digital-transformation-of-smes_bdb9256a-en.html.

[29] L. Oldemeyer, A. Jede, and F. Teuteberg. «Investigation of Artificial Intelligence in SMEs: A Systematic Review of the State of the Art and the Main Implementation Challenges». In: *Management Review Quarterly* (2024). DOI: 10.1007/s11301-024-00405-4.

[30] J. J. Daspit, J. J. Chrisman, T. Ashton, and N. Evangelopoulos. «Family Firm Heterogeneity: A Definition, Common Themes, Scholarly Progress, and Directions Forward». In: *Family Business Review* 34.3 (2021), pp. 296–322. DOI: 10.1177/08944865211008350.

[31] Eurostat. *E-business Integration.* Statistics Explained, European Commission, Luxembourg. Data extracted in May 2024. 2025. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=E-business_integration.

[32] Eurostat. *Digital Economy and Society Statistics – Enterprises.* Statistics Explained, European Commission, Luxembourg. 2024. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_enterprises.

[33] KfW Research. *SMEs Spend Seven per Cent of Their Working Time on Administrative Processes.* Tech. rep. KfW Group, 2025. URL: https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Fokus-Volkswirtschaft/Fokus-englische-Dateien/Fokus-2025-EN/Focus-No.-495-April-2025-Bureaucracy.pdf.

[34] Eurostat. *8% of EU Enterprises Used AI Technologies in 2023.* Eurostat News Articles. May 2024. URL: https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20240529-2.

[35] OECD. *SME Digitalisation in 2024 – Managing Shocks and Transitions*. Tech. rep. OECD SME and Entrepreneurship Papers, OECD Publishing, Paris, 2024. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/09/sme-digitalisation-to-manage-shocks-and-transitions_735fc44d/eb4ec9ac-en.pdf.

[36] OECD. *AI Adoption by Small and Medium-Sized Enterprises*. Tech. rep. OECD Secretariat for the G7 Industry, Digital and Technology Ministerial, OECD Publishing, Paris, 2025. URL: https://www.oecd.org/en/publications/ai-adoption-by-small-and-medium-sized-enterprises_426399c1-en.html.

[37] R. Press. *AI in Business: Small Firms Closing In*. Tech. rep. Office of Advocacy, U.S. Small Business Administration, 2025. URL: https://advocacy.sba.gov/wp-content/uploads/2025/09/Research-Spotlight-AI-in-Business-Small-Firms-Closing-In_-092425.pdf.

[38] S. S. Muhammad, B. L. Dey, M. M. Kamal, L. Samuel, and E. A. Alzeiby. «Digital Transformation or Digital Divide? SMEs' Use of AI During Global Crisis». In: *Technological Forecasting & Social Change* (2025).

[39] OECD. *Generative AI and the SME Workforce*. Tech. rep. OECD SME and Entrepreneurship Papers, OECD Publishing, Paris, 2025. URL: https://www.oecd.org/en/publications/generative-ai-and-the-sme-workforce_2d08b99d-en.html.

[40] K. Rajaram and P. N. Tinguely. «Generative Artificial Intelligence in Small and Medium Enterprises: Navigating Its Promises and Challenges». In: *Business Horizons* 67.5 (2024), pp. 629–648. DOI: 10.1016/j.bushor.2024.06.002.

[41] D. Horváth and R. Z. Szabó. «Driving Forces and Barriers of Industry 4.0: Do Multinational and Small and Medium-Sized Companies Have Equal Opportunities?» In: *Technological Forecasting and Social Change* 146 (2019), pp. 119–132. DOI: 10.1016/j.techfore.2019.05.021.

[42] L. S. Závodná, M. Überwimmer, and E. Frankus. «Barriers to the Implementation of Artificial Intelligence in Small and Medium-Sized Enterprises: Pilot Study». In: *Journal of Economics and Management (JEM)* (2024). DOI: 10.22367/jem.2024.46.13.

[43] McKinsey & Company. *The State of AI: Global Survey 2025*. Tech. rep. QuantumBlack Insights, 2025. URL: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai.

[44] O. Sysoiev. «Dynamics of LLM Inference Costs and AI Startup Growth: An Empirical Analysis». In: *Nauki Ekonomiczne (Akademia Piotrkowska)* (2025). URL: https://czasopismanaukowe.mazowiecka.edu.pl/index.php/ne/article/download/1108/1072.

[45] S. Feuerriegel, J. Hartmann, C. Janiesch, and P. Zschech. «Generative AI». In: *Business & Information Systems Engineering* (2023). DOI: 10.1007/s12599-023-00834-7.

[46] Gartner. *Gartner Predicts More than 30% of the Increase in Demand for APIs Will Come from AI and Tools Using LLMs by 2026.* Press release, Gartner Newsroom. 2024. URL: https://www.gartner.com/en/newsroom/press-releases/2024-03-20-gartner-predicts-more-than-30-percent-of-the-increase-in-demand-for-apis-will-come-from-ai-and-tools-using-llms-by-2026.

[47] M. Zhang, J. Cao, X. Shen, and Z. Cui. «EdgeShard: Efficient LLM Inference via Collaborative Edge Computing». In: *IEEE Internet of Things Journal* (2024). DOI: 10.1109/jiot.2024.3524255.

[48] J. Pereira, J. M. López, and X. Garmendia. «Leveraging Open-Source LLMs for Software Engineering Education and Training». In: *IEEE CSEET 2024*. 2024. DOI: 10.1109/cseet62301.2024.10663055.

[49] D. Hamdhana. «A Low-Code Approach to Q&A on Care Records Using Flowise AI with LLM Integration and RAG Method». In: *Jurnal Ilmiah Penelitian dan Pembelajaran Informatika (JIPI)* (2024). DOI: 10.29100/jipi.v9i4.6978.

[50] Ernst & Young. *Low-Code No-Code Platforms and a Culture of Innovation.* EY Alliances, US. 2024. URL: https://www.ey.com/en_us/alliances/low-code-no-code-platforms-and-a-culture-of-innovation.

[51] L. Berberi, V. Kozlov, G. Nguyen, J. Sáinz-Pardo Díaz, A. Calatrava, G. Moltó, V. Tran, and Á. López García. «Machine Learning Operations Landscape: Platforms and Tools». In: *Artificial Intelligence Review* 58 (2025), Article 167. DOI: 10.1007/s10462-025-11164-3.

[52] Microsoft and LinkedIn. *2024 Work Trend Index Annual Report: AI at Work Is Here. Now Comes the Hard Part.* Tech. rep. Microsoft Corporation, 2024. URL: https://news.microsoft.com/source/2024/05/08/microsoft-and-linkedin-release-the-2024-work-trend-index-on-the-state-of-ai-at-work/.

[53] D. F. Engstrom and J. B. Gelbach. «Legal Tech, Civil Procedure, and the Future of Adversarialism». In: *University of Pennsylvania Law Review* 169 (2021), pp. 1001–1100. URL: https://scholarship.law.upenn.edu/penn_law_review/vol169/iss4/2/.

[54] Boston Consulting Group. *AI and the Future of Software: BCG Executive Perspectives, April 2026.* Tech. rep. Boston Consulting Group, 2026. URL: https://www.bcg.com/assets/2026/executive-perspectives-ai-and-the-future-of-software.pdf.

[55] Gartner. *Gartner Predicts Embedded AI in Cloud ERP Applications Will Drive a 30 Percent Faster Financial Close by 2028.* Press release, Gartner Newsroom. 2026. URL: https://www.gartner.com/en/newsroom/press-releases/2026-02-24-gartner-predicts-embedded-ai-in-cloud-erp-applications-will-drive-a-30-percent-faster-financial-close-by-2028.

[56] S. Thota. «A Secure Multi-Tenant AI Framework for Enterprise CRM Automation on Salesforce Cloud Platforms». In: *International Journal of Engineering Technology and Computer Science (IJETCSIT)* (2025). DOI: 10.63282/3050-9246.ijetcsit-v6i2p113.

[57] F. Bäumer, S. Denisov, and B. Sirvend. «Tackling the 'We Have No Data' Challenge: Domain-Specific Machine Translation in SMEs». In: *PATTERNS 2022, International Conference on Pervasive Patterns and Applications*. 2022. URL: https://personales.upv.es/thinkmind/dl/conferences/patterns/patterns_2022/patterns_2022_1_10_78001.pdf.

[58] L. Eberhardsteiner. «Adopting AI in SMEs». MA thesis. Johannes Kepler Universität Linz, 2024. URL: https://epub.jku.at/obvulihs/download/pdf/11427103.

[59] B. Florina, D. A. Bodislav, and A. M. Călin. «Empowering SMEs through Generative AI: Opportunities, Challenges, and Strategic Implications for Sustainable Innovation». In: *European Journal of Sustainable Development (EJSD)* 14.4 (2025). DOI: 10.14207/ejsd.2025.v14n4p27.

[60] World Bank Group. *Digital Progress and Trends Report 2025: Strengthening AI for Inclusive Development*. Tech. rep. World Bank Publications, Washington, DC, 2025. URL: https://www.worldbank.org/en/publication/dptr2025-ai-foundations.

[61] OECD. «Implementation Challenges that Hinder the Strategic Use of AI in Government». In: *Governing with Artificial Intelligence*. OECD Publishing, Paris, 2025. URL: https://www.oecd.org/en/publications/2025/06/governing-with-artificial-intelligence_398fa287/full-report/implementation-challenges-that-hinder-the-strategic-use-of-ai-in-government_05cfe2bb.html.

[62] McKinsey & Company. *LLM to ROI: How to Scale Gen AI in Retail*. Tech. rep. McKinsey Industries: Retail, 2024. URL: https://www.mckinsey.com/industries/retail/our-insights/llm-to-roi-how-to-scale-gen-ai-in-retail.

[63] E. Brynjolfsson, D. Rock, and C. Syverson. «The Productivity J-Curve: How Intangibles Complement General Purpose Technologies». In: *American Economic Journal: Macroeconomics* 13.1 (2021), pp. 333–372. DOI: 10.1257/mac.20180386.

[64] Microsoft. *Microsoft 365 Copilot Architecture and How It Works*. Microsoft Learn. 2024. URL: https://learn.microsoft.com/en-us/microsoft-365/copilot/microsoft-365-copilot-architecture.

[65] OECD. *Generative AI and the SME Workforce: New Survey Evidence*. Tech. rep. OECD Publishing, Paris, 2025. DOI: 10.1787/2d08b99d-en.

[66] L. Ardito, R. Filieri, E. Raguseo, and C. Vitari. «Artificial Intelligence Adoption and Revenue Growth in European SMEs: Synergies with IoT and Big Data Analytics». In: *Internet Research* 35.4 (2025), pp. 1508–1534. DOI: 10.1108/INTR-02-2024-0195.

[67] G7 Industry, Digital and Technology Ministers. *The SME AI Adoption Blueprint*. Tech. rep. G7 Canada Presidency / OECD, 2025. URL: https://ised-isde.canada.ca/site/ised/en/sme-ai-adoption-blueprint.

[68] A. Katsinis, J. Lagüera-González, L. Di Bella, L. Odenthal, M. Hell, and B. Lozar. *Annual Report on European SMEs 2023/2024: SME Performance Review*. Tech. rep. Joint Research Centre, European Commission, 2024. DOI: 10.2826/355464.

[69] NASSCOM and Ernst & Young. *NASSCOM AI Adoption Index: Tracking India's Sectoral Progress on AI Adoption.* Tech. rep. NASSCOM, New Delhi, 2020.

[70] Deloitte AI Institute. *The State of AI in the Enterprise, 2026: The Untapped Edge.* Tech. rep. Deloitte, 2026. URL: https://www.deloitte.com/us/en/what-we-do/capabilities/applied-artificial-intelligence/content/state-of-ai-in-the-enterprise.html.

[71] EIT Digital. *Creation of a Taxonomy for the European AI Ecosystem.* Tech. rep. Cross-KIC Activity, European Institute of Innovation and Technology, 2020.

[72] R. Schenker, V. Faidherbe, K. Mediratta, and L. Than. *Why Agentic AI Demands a New Architecture.* Tech. rep. Bain & Company, 2026. URL: https://www.bain.com/insights/why-agentic-ai-demands-a-new-architecture/.

[73] M. Bianchini and M. Lasheras Sancho. *SME Digitalisation for Competitiveness: The 2025 OECD D4SME Survey.* Tech. rep. OECD SME and Entrepreneurship Papers, OECD Publishing, Paris, 2025.

[74] N. Omrani, N. Rejeb, A. Maalaoui, M. Dabić, and S. Kraus. «Drivers of Digital Transformation in SMEs». In: *IEEE Transactions on Engineering Management* (2022). DOI: 10.1109/TEM.2022.3215727.

[75] S. Ramachandran, P. Himpens, and L. Janssens. *What SMEs Really Want: How Banks Can Successfully Meet Evolving Needs.* Tech. rep. Boston Consulting Group, 2024.

[76] M. Paul. «Organisational Readiness for Artificial Intelligence Adoption». MA thesis. University of Turku, 2020. URL: https://www.utupub.fi/handle/10024/150266.

[77] Eurostat. *The Use of Artificial Intelligence (AI) Technologies in the European Union — Key Results — 2026 Edition.* Tech. rep. Publications Office of the European Union, Luxembourg, 2026. DOI: 10.2785/9221093. URL: https://ec.europa.eu/eurostat/documents/7870049/23260410/KS-01-26-009-EN-N.pdf.

[78] Accenture. *The Art of AI Maturity: Advancing from Practice to Performance.* Tech. rep. Accenture, 2021. URL: https://www.accenture.com/us-en/insights/artificial-intelligence/ai-maturity-and-transformation.

[79] A. Salmen. «Employing RPA and AI to Automize Order Entry Process with Individual and Small-Sized Structures: A SME Business Case Study». In: *Acta Academica Karviniensia* XXII.2 (2022), pp. 78–96. DOI: 10.25142/aak.2022.017.

[80] World Economic Forum and IndiaAI. *Transforming Small Businesses: An AI Playbook for India's SMEs.* Tech. rep. World Economic Forum, in collaboration with IndiaAI and BCG X, 2025. URL: https://reports.weforum.org/docs/WEF_Transforming_Small_Businesses_2025.pdf.

[81] A. O. Ikumoro and M. S. Jawad. «Intention to Use Intelligent Conversational Agents in e-Commerce among Malaysian SMEs». In: *International Journal of Academic Research in Business and Social Sciences* 9.11 (2019), pp. 205–235. DOI: 10.6007/IJARBSS/v9-i11/6544.

[82] W. F. Ridho. «An Examination of the Opportunities and Challenges of Conversational Artificial Intelligence in Small and Medium Enterprises». In: *Review of Business and Economics Studies* 11.3 (2023), pp. 6–17. DOI: 10.26794/2308-944X-2023-11-3-6-17.

[83] B. O. Eboseremen, A. O. Ogedengbe, et al. «Developing an AI-Driven Personalization Pipeline for Customer Retention in Investment Platforms». In: *Journal of Frontiers in Multidisciplinary Research* 3.1 (2022).

[84] E. Karakurt and A. Akbulut. «Retrieval-Augmented Generation (RAG) and Large Language Models (LLMs) for Enterprise Knowledge Management and Document Automation: A Systematic Literature Review». In: *Applied Sciences* 16.1 (2025), p. 368. DOI: 10.3390/app16010368.

[85] World Economic Forum. *Empowering Small and Medium Enterprises (SMEs) for Innovation and Success*. Tech. rep. World Economic Forum, Geneva, 2024.

[86] Various Authors. «Application of Retrieval-Augmented Generation for Interactive Industrial Knowledge Management via a Large Language Model». In: *Computers in Industry* (2025). DOI: 10.1016/j.compind.2025.104248.

[87] Deloitte Netherlands. *Knowledge-Enriched Agentic AI Workflows: Intelligent and Responsible Task Execution for Business Growth*. Tech. rep. Deloitte, 2024.

[88] I. Barberá. *AI Privacy Risks and Mitigations: Large Language Models (LLMs)*. Tech. rep. European Data Protection Board, Support Pool of Experts Programme, 2025.

[89] M. S. Soudi and M. Bauters. «AI Guidelines and Ethical Readiness Inside SMEs: A Review and Recommendations». In: *Digital Society* 3.3 (2024). DOI: 10.1007/s44206-024-00087-1.

[90] Y. Zhao et al. «HoH: A Dynamic Benchmark for Evaluating the Impact of Outdated Information on Retrieval-Augmented Generation». In: *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025)*. 2025. URL: https://aclanthology.org/2025.acl-long.301/.

[91] A. Agarwal, C. Singhal, and R. Thomas. *AI-Powered Decision Making for the Bank of the Future*. Tech. rep. McKinsey & Company, Global Banking & Securities, 2021. URL: https://www.mckinsey.com/industries/financial-services/our-insights/ai-powered-decision-making-for-the-bank-of-the-future.

[92] N. C. Igbokwe, C. O. Nwamekwe, P. S. Aguh, and O. C. Ezeanyim. «A Comparative Analysis of Machine Learning Models for Inventory Demand Forecasting in a Food Manufacturing SME». In: *International Journal of Information Science and Knowledge* 2.3 (2025), pp. 35–48.

[93] L. Griesch, J. Rittelmeyer, and K. Sandkuhl. «Towards AI as a Service for Small and Medium-Sized Enterprises (SME)». In: *The Practice of Enterprise Modeling. PoEM 2023*. Vol. 497. Lecture Notes in Business Information Processing. Springer, Cham, 2024. DOI: 10.1007/978-3-031-48583-1_3.

[94] Menlo Ventures. *2025: The State of Generative AI in the Enterprise*. Tech. rep. Survey of 495 U.S. enterprise AI decision-makers, Nov 2025. Menlo Ventures, 2025. URL: https://menlovc.com/perspective/2025-the-state-of-generative-ai-in-the-enterprise/.

[95] Eurostat. *Digitalisation in Europe — 2025 Edition*. Tech. rep. Publications Office of the European Union, Luxembourg, 2025. DOI: 10.2785/3102705.

[96] S. Afrin, S. Roksana, and R. Akram. «AI-Enhanced Robotic Process Automation: A Review of Intelligent Automation Innovations». In: *IEEE Access* 13 (2025), pp. 173–197. DOI: 10.1109/ACCESS.2024.3513279.

[97] L. Patricio, L. Varela, and Z. Silveira. «Integration of Artificial Intelligence and Robotic Process Automation: Literature Review and Proposal for a Sustainable Model». In: *Applied Sciences* 14.21 (2024), p. 9648. DOI: 10.3390/app14219648.

[98] Grand View Research. *Intelligent Document Processing Market Size, Share and Trends Analysis Report, 2025–2030*. Tech. rep. Report ID: GVR-4-68038-063-2. Grand View Research, 2024.

[99] Ardent Partners. *Accounts Payable Metrics That Matter in 2025*. Tech. rep. Survey of 212 AP professionals. Ardent Partners, 2025.

[100] Wolters Kluwer. *2026 Future Ready Lawyer Survey: Building Confidence in an AI Era*. Tech. rep. Survey of 810 lawyers, August 2025. Wolters Kluwer Legal & Regulatory, 2026.

[101] Everest Group. *Intelligent Document Processing (IDP) Products PEAK Matrix Assessment 2024*. Tech. rep. Everest Group, 2024. URL: https://www.abbyy.com/resources/report/intelligent-document-processing-everest-group-2024/.

[102] M. Arcaro and A. Machado. *IDC MarketScape: Worldwide Intelligent Document Processing Software 2023–2024 Vendor Assessment*. Tech. rep. US49988723. International Data Corporation (IDC), 2023. URL: https://www.idc.com/getdoc.jsp?containerId=US49988723.

[103] F. F. Eiffe, S. Riso, E. Biaggi, J. Van Loo, and G. Miliadis. *SME Digitalisation in the EU: Trends, Policies and Impacts*. Tech. rep. Eurofound and Cedefop, Publications Office of the European Union, Luxembourg, 2025. URL: https://www.eurofound.europa.eu/en/publications/2025/sme-digitalisation-eu-trends-policies-and-impacts.

[104] K. Marcinkova, A. J. Sujova, and R. Durica. «Implementing AI Chatbots in Customer Service Optimization — A Case Study in Micro-Enterprise». In: *Information* 16.12 (2025), p. 1078. DOI: 10.3390/info16121078.

[105] W. Cui, D. Yu, X. Jiao, Z. Meng, G. Zhang, Q. Wang, S. Y. Guo, and I. King. «Recent Advances in Speech Language Models: A Survey». In: *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025)*. 2025, pp. 13943–13970. DOI: 10.18653/v1/2025.acl-long.682.

[106] T. Le Dinh, M.-C. Vu, and G. T. C. Tran. «Artificial Intelligence in SMEs: Enhancing Business Functions Through Technologies and Applications». In: *Information* 16.5 (2025), p. 415. DOI: 10.3390/info16050415.

[107] S. Bonechi, G. Palma, M. Caronna, M. Ugolini, A. Massaro, and A. Rizzo. «Enhancing Customer Support in Banking: Leveraging AI for Efficient Ticket Classification». In: *Procedia Computer Science*. Vol. 246. 2024, pp. 128–137. DOI: 10.1016/j.procs.2024.09.235.

[108] Gartner. *Gartner Predicts Conversational AI Will Reduce Contact Center Agent Labor Costs by $80 Billion in 2026*. Press release. 2022. URL: https://www.gartner.com/en/newsroom/press-releases/2022-08-31-gartner-predicts-conversational-ai-will-reduce-contac.

[109] Gartner. *Gartner Predicts Agentic AI Will Autonomously Resolve 80% of Common Customer Service Issues by 2029*. Press release. 2025. URL: https://www.gartner.com/en/newsroom/press-releases/2025-03-05-gartner-predicts-agentic-ai-will-autonomously-resolve-80-percent-of-common-customer-service-issues-without-human-intervention-by-20290.

[110] Salesforce. *State of Service Report, Seventh Edition*. Tech. rep. Survey of 6,500 service professionals, April–June 2025. Salesforce Research, 2025.

[111] Grand View Research. *AI Voice Agents Market Size, Share and Trends Analysis Report, 2026–2033*. Tech. rep. Grand View Research, 2025.

[112] A. J. Mou. «MarTech Stack Adoption in SMEs: A Review of Automation, CRM, and AI Integration». In: *American Journal of Advanced Technology and Engineering Solutions* 1.1 (2025), pp. 348–381. DOI: 10.63125/y8j1zh51.

[113] L. González-Flores, J. Rubiano-Moreno, and G. Sosa-Gómez. «The Relevance of Lead Prioritization: A B2B Lead Scoring Model Based on Machine Learning». In: *Frontiers in Artificial Intelligence* 8 (2025), p. 1554325. DOI: 10.3389/frai.2025.1554325.

[114] I. K. Magableh, M. H. Mahrouq, M. A. Ta'Amnha, and H. A. Riyadh. «The Role of Marketing Artificial Intelligence in Enhancing Sustainable Financial Performance of Medium-Sized Enterprises». In: *Sustainability* 16.24 (2024), p. 11279. DOI: 10.3390/su162411279.

[115] McKinsey & Company. *The Value of Getting Personalization Right—or Wrong—Is Multiplying*. Tech. rep. McKinsey & Company, 2021. URL: https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-value-of-getting-personalization-right-or-wrong-is-multiplying.

[116] HubSpot. *State of Marketing Report, Tenth Edition*. Tech. rep. Survey of over 1,500 global marketers, 2025. HubSpot, 2026.

[117] J. J. Pan, J. Wang, and G. Li. «Survey of Vector Database Management Systems». In: *The VLDB Journal* 33.5 (2024), pp. 1591–1615. DOI: 10.1007/s00778-024-00864-x.

[118] MarketsandMarkets. *Vector Database Market — Global Forecast to 2030.* Tech. rep. TC 8828. MarketsandMarkets Research Private Ltd., 2025. URL: https://www.marketsandmarkets.com/Market-Reports/vector-database-market-112683895.html.

[119] L.-C. Chen, M. S. Pardeshi, Y.-X. Liao, and K.-C. Pai. «Application of Retrieval-Augmented Generation for Interactive Industrial Knowledge Management via a Large Language Model». In: *Computer Standards & Interfaces* 94 (2025), p. 103995. DOI: 10.1016/j.csi.2025.103995.

[120] S. Nishisako, T. Higashi, and F. Wakao. «Reducing Hallucinations and Trade-Offs in Responses in Generative AI Chatbots for Cancer Information: Development and Evaluation Study». In: *JMIR Cancer* 11.1 (2025), e70176. DOI: 10.2196/70176.

[121] S. Xu, Z. Yan, C. Dai, and F. Wu. «MEGA-RAG: A Retrieval-Augmented Generation Framework with Multi-Evidence Guided Answer Refinement for Mitigating Hallucinations of LLMs in Public Health». In: *Frontiers in Public Health* 13 (2025), p. 1635381. DOI: 10.3389/fpubh.2025.1635381.

[122] T. Gelashvili-Luik, P. Vihma, and I. Pappel. «Navigating the AI Revolution: Challenges and Opportunities for Integrating Emerging Technologies into Knowledge Management Systems. Systematic Literature Review». In: *Frontiers in Artificial Intelligence* 8 (2025), p. 1595930. DOI: 10.3389/frai.2025.1595930.

[123] M. Chui, J. Manyika, J. Bughin, R. Dobbs, C. Roxburgh, H. Sarrazin, G. Sands, and M. Westergren. *The Social Economy: Unlocking Value and Productivity Through Social Technologies.* Tech. rep. McKinsey Global Institute, 2012.

[124] Mordor Intelligence. *Enterprise Search Market — Growth, Trends, and Forecast (2026–2031).* Tech. rep. Mordor Intelligence, 2026.

[125] Gartner. *Gartner Predicts 75% of Analytics Content to Use GenAI for Enhanced Contextual Intelligence by 2027.* Press release. Survey of 403 analytics/AI leaders, Oct–Dec 2024. 2025. URL: https://www.gartner.com/en/newsroom/press-releases/2025-06-18-gartner-predicts-75-percent-of-analytics-content-to-use-genai-for-enhanced-contextual-intelligence-by-2027.

[126] G. Naren Shailesh, M. Pavithran, A. Rahul Hari Venkat, and P. Kaliappan. «Conversational BI: Natural Language Interface to Business Dashboards». In: *International Journal of Engineering Research & Technology* 14.12 (2025). DOI: 10.17577/IJERTV14IS120229.

[127] Z. Yuan, Y. Wang, T. Sun, and H. Xu. «Democratizing Data Analytics Products to Drive SME Growth: A Natural Experiment». In: *Manufacturing & Service Operations Management* 27.4 (2025), pp. 1053–1067. DOI: 10.1287/msom.2024.1138.

[128] H. E. Sanches, A. T. Possebom, and L. B. R. Aylon. «Churn Prediction for SaaS Company with Machine Learning». In: *Innovation & Management Review* 22.2 (2025), pp. 130–142. DOI: 10.1108/INMR-06-2023-0101.

[129] OECD. *Empowering SMEs in the Age of AI: The 2025 OECD D4SME Survey.* Tech. rep. OECD Publishing, Paris, 2026. DOI: 10.1787/bf5a9816-en.

[130] J. S. Allen. «Monitoring AI Adoption in the US Economy». In: *FEDS Notes* (2026). DOI: 10.17016/2380-7172.4032.

[131] Gartner. *Gartner Predicts 40% of Enterprise Apps Will Feature Task-Specific AI Agents by 2026.* Press release. 2025. URL: https://www.gartner.com/en/newsroom/press-releases/2025-08-26-gartner-predicts-40-percent-of-enterprise-apps-will-feature-task-specific-ai-agents-by-2026-up-from-less-than-5-percent-in-2025.

[132] A. Neri. *AI's Future: Plotting a Path to Competitiveness and Digital Sovereignty.* Tech. rep. World Economic Forum, 2026. URL: https://www.weforum.org/stories/2026/01/ai-s-distributed-future-a-new-path-to-competitiveness-and-digital-sovereignty/.

[133] Flexera. *2024 State of the Cloud Report.* Tech. rep. Flexera, 2024. URL: https://www.flexera.com/blog/finops/cloud-computing-trends-flexera-2024-state-of-the-cloud-report/.

[134] L. Oldemeyer, A. Jede, and F. Teuteberg. «Unlocking AI Potential in SMEs: Applications, Benefits, and Preliminary Considerations for Implementation». In: *Journal of the Knowledge Economy* (2026). DOI: 10.1007/s13132-026-03242-w.

[135] J. Schwaeke, A. Peters, D. K. Kanbach, S. Kraus, and P. Jones. «The New Normal: The Status Quo of AI Adoption in SMEs». In: *Journal of Small Business Management* 63.3 (2024), pp. 1297–1331. DOI: 10.1080/00472778.2024.2379999.

[136] E. M. Rogers. *Diffusion of Innovations.* 5th. New York: Free Press, 2003.

[137] I. Sahin. «Detailed Review of Rogers' Diffusion of Innovations Theory and Educational Technology-Related Studies Based on Rogers' Theory». In: *The Turkish Online Journal of Educational Technology* 5.2 (2006).

[138] R. G. Cooper and S. J. Edgett. «Stage-Gate and the Critical Success Factors for New Product Development». In: *BPTrends* (2006). URL: https://www.bptrends.info/wp-content/publicationfiles/07-06-ART-Stage-GateForProductDev-Cooper-Edgett1.pdf.

[139] L. G. Tornatzky and M. Fleischer. *The Processes of Technological Innovation.* Lexington, MA: Lexington Books, 1990.

[140] T. Oliveira and M. F. Martins. «Literature Review of Information Technology Adoption Models at Firm Level». In: *The Electronic Journal Information Systems Evaluation* 14.1 (2011), pp. 110–121.

[141] StepUp StartUps Consortium. *Unlocking the Potential of AI in Business Support Processes.* Tech. rep. European Commission, DG CNECT, 2025. URL: https://digital-strategy.ec.europa.eu/en/library/unlocking-potential-ai-business-support-processes.

[142] European Commission, Joint Research Centre. *Digital Maturity Assessment Questionnaire for SMEs.* Tech. rep. European Commission, JRC and European Digital Innovation Hubs Network, 2023. URL: https://european-digital-innovation-hubs.ec.europa.eu/.

[143] S. G. Ayinaddis. «Artificial Intelligence Adoption Dynamics and Knowledge in SMEs and Large Firms: A Systematic Review and Bibliometric Analysis». In: *Journal of Innovation and Knowledge* 10.3 (2025), p. 100682. DOI: 10.1016/j.jik.2025.100682.

[144] L. M. Ellram. «Total Cost of Ownership: An Analysis Approach for Purchasing». In: *International Journal of Physical Distribution and Logistics Management* 25.8 (1995), pp. 4–23. DOI: 10.1108/09600039510099928.

[145] McKinsey and Company. *Moving Past Gen AI's Honeymoon Phase: Seven Hard Truths for CIOs to Get from Pilot to Scale*. Tech. rep. McKinsey and Company, 2024. URL: https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/moving-past-gen-ais-honeymoon-phase-seven-hard-truths-for-cios-to-get-from-pilot-to-scale.

[146] N. de Bellefonds, T. Charanya, J. Apotheker, and P. Forth. *Where's the Value in AI?* Tech. rep. Boston Consulting Group, 2024. URL: https://www.bcg.com/publications/2024/wheres-value-in-ai.

[147] Gartner. *Gartner Predicts 30% of Generative AI Projects Will Be Abandoned After Proof of Concept by End of 2025*. Press release, Gartner Newsroom. 2024. URL: https://www.gartner.com/en/newsroom/press-releases/2024-07-29-gartner-predicts-30-percent-of-generative-ai-projects-will-be-abandoned-after-proof-of-concept-by-end-of-2025.

[148] Cedefop. *Skills Empower Workers in the AI Revolution: First Findings from Cedefop's AI Skills Survey*. Tech. rep. Cedefop, 2025. URL: https://www.cedefop.europa.eu/en/publications/9201.

[149] K. Lazaros, A. G. Vrahatis, and S. Kotsiantis. «Human-in-the-Loop Artificial Intelligence: A Systematic Review of Concepts, Methods, and Applications». In: *Entropy* 28.4 (2026), p. 377. DOI: 10.3390/e28040377.

[150] National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. Tech. rep. NIST AI 100-1. National Institute of Standards and Technology, 2023. DOI: 10.6028/NIST.AI.100-1.

[151] Qualtrics. *2026 Consumer Experience Trends Report*. Tech. rep. Qualtrics XM Institute, 2025. URL: https://www.qualtrics.com/articles/news/ai-powered-customer-service-fails-at-four-times-the-rate-of-other-tasks/.

[152] C. Ferraro, V. Demsar, S. Sands, M. Restrepo, and C. Campbell. «The Paradoxes of Generative AI-Enabled Customer Service: A Guide for Managers». In: *Business Horizons* 67.5 (2024), pp. 549–559. DOI: 10.1016/j.bushor.2024.04.013.

[153] S. A. Kalleparambil and M. Akoum. «AI Organisational Readiness for SMEs: A Tailored Model for AI Adoption Success». In: *Journal of Information and Knowledge Management* 24.06 (2025). DOI: 10.1142/S0219649225500765.

[154] J. Opara-Martins, R. Sahandi, and F. Tian. «Critical Review of Vendor Lock-in and Its Impact on Adoption of Cloud Computing». In: *International Conference on Information Society (i-Society 2014)*. IEEE, 2014. URL: https://eprints.bournemouth.ac.uk/22467/.

[155] D. Sculley, G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young, J.-F. Crespo, and D. Dennison. «Hidden Technical Debt in Machine Learning Systems». In: *Advances in Neural Information Processing Systems 28 (NeurIPS)*. 2015. URL: https://proceedings.neurips.cc/paper_files/paper/2015/hash/86df7dcfd896fcaf2674f757a2463eba-Abstract.html.

[156] R. M. Rasdi and N. Umar Baki. «Navigating the AI Landscape in SMEs: Overcoming Internal Challenges and External Obstacles for Effective Integration». In: *PLOS One* (2025), e0323249. DOI: 10.1371/journal.pone.0323249.

[157] A. Asatiani and E. Penttinen. «Turning Robotic Process Automation into Commercial Success – Case OpusCapita». In: *Journal of Information Technology Teaching Cases* 6.2 (2016), pp. 67–74. DOI: 10.1057/jittc.2016.5.

[158] M. Lacity, L. P. Willcocks, and A. Craig. *Robotic Process Automation at Telefónica O2*. Working Paper 15/02. The Outsourcing Unit, London School of Economics, 2015. URL: https://eprints.lse.ac.uk/64516/.

[159] F. F. Xu et al. «TheAgentCompany: Benchmarking LLM Agents on Consequential Real World Tasks». In: *Advances in Neural Information Processing Systems (NeurIPS) 38, Datasets and Benchmarks Track*. 2025.

[160] J. Ryseff, B. De Bruhl, and S. J. Newberry. *The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed*. Tech. rep. RR-A2680-1. RAND Corporation, 2024. DOI: 10.7249/RRA2680-1.