IIBM INSTITUTE OF BUSINESS MANAGEMENT

Doctor of Business Administration

DOCTORAL DISSERTATION

Monetizing AI in SaaS: An Empirical Study of Pricing and Packaging Models for AI Features.

Hitesh Sarda

A Dissertation

Submitted to the IIBM Faculty/Dean Dr. Latika V Shanbhag

for a DBA degree:

Doctor of Business Administration

Swiss International Management Academy, Switzerland

March 2026

DECLARATION

I declare that I have elaborated the doctoral dissertation entitled

"Monetizing AI in SaaS: An Empirical Study of Pricing and Packaging Models for AI Features."

independently, under the expert supervision of my thesis/dissertation supervisor, and using only the literature and sources cited therein, and that the thesis/dissertation was not used within the scope of a different university programme of study or to obtain the same degree or a different degree. I consent to the publication of the thesis/dissertation in accordance with legislation and with IIBM Institute internal regulations.



SIMA, Switzerland

Hitesh Sarda

7th March 2026

ACKNOWLEDGEMENTS

This dissertation marks a significant milestone in my academic and professional journey, and it would not have been possible without the unwavering support, guidance, and encouragement of several individuals.

First and foremost, I extend my heartfelt gratitude to my research guide, **Dr. Latika Shanbhag**, for her patient mentorship, insightful feedback, and consistent encouragement throughout this research process. Her experience, dedication, and clarity helped shape this work with both academic rigor and practical relevance.

I am deeply thankful to my **family**, whose unconditional love, support, and understanding gave me the strength and motivation to pursue this demanding journey. Their patience during long evenings, missed weekends, and countless moments of quiet sacrifice has meant the world to me.

I would also like to express my appreciation to the entire academic and administrative teams at the **Indian Institute of Business Management (IIBM)** and the **Swiss International Management Academy (SIMA)**. Their professionalism, responsiveness, and commitment to academic excellence provided a strong foundation for this dissertation and my overall learning experience.
This work reflects the collective support of all these individuals and institutions. To each of you, I say thank you.

ABSTRACT

The rapid integration of Artificial Intelligence (AI) into Software as a Service (SaaS) products represents a fundamental structural disruption to established software business models. Unlike traditional SaaS architectures premised on high fixed development costs and near-zero marginal distribution expenses, AI-powered features introduce significant, variable computational costs that scale linearly with usage. This structural break renders conventional per-seat pricing frameworks inadequate and creates a critical misalignment between vendor cost structures and customer demands for budget predictability. Despite the strategic urgency of this challenge, existing academic literature remains fragmented, largely descriptive, and empirically underdeveloped. This study addresses that gap through a rigorous mixed-methods empirical investigation into the pricing and packaging strategies adopted by B2B SaaS firms generating between $1 million and $100 million in Annual Recurring Revenue (ARR).

The research employed a sequential mixed-methods design comprising two integrated phases. The quantitative phase involved systematic classification of public-facing pricing pages across 371 mid-sized SaaS companies, mapping the market distribution of subscription, usage-based, credit-based, and hybrid monetization models. The qualitative phase comprised semi-structured interviews with senior SaaS executives, conducted to surface the strategic rationales, operational constraints, and competitive pressures that underpin observed pricing decisions. Thematic analysis was applied to interview data following the framework established by Braun and Clarke (2006), and findings were integrated through a convergent interpretive synthesis.

The results reveal a monetization landscape characterised by significant heterogeneity and ongoing experimentation. Pure subscription models remain overwhelmingly dominant, observed in 191 of the 371 firms analysed, while a measurable structural transition toward hybrid architectures is evident, particularly among firms in the $50–$100 million ARR tier. Qualitative findings identified three primary operational tensions driving these configurations: the imperative to protect gross margins against unpredictable AI inference costs; widespread enterprise resistance to pure consumption-based billing; and competitive anchoring driven by the ubiquity of low-cost consumer AI tools. Credit-based systems emerged as the dominant abstraction mechanism, enabling firms to translate variable token costs into psychologically predictable pricing units acceptable to enterprise buyers.

These findings carry significant theoretical and managerial implications. Theoretically, the study extends SaaS pricing literature into variable-cost environments, empirically validates hybrid monetization as a pragmatic transitional equilibrium, and enriches AI business model transformation theory by grounding abstract conceptual claims in observable commercial realities. Managerially, the research yields a stage-based strategic framework linking monetization architecture to firm maturity, and provides actionable guidance on margin management, overage policy design, and value communication. The study concludes that the monetization of AI is not a tactical pricing adjustment but a fundamental redesign of the SaaS business model, requiring organizational agility to continuously align dynamic cost structures with sustainable value capture.

Table of Contents

Chapter 1. Introduction 9
1.1 Background 9
1.2 Problem Statement 10
1.3 Objectives 11
1.3.1 Primary Objective 12
1.3.2 Secondary Objectives 12
1.4 Significance 12
1.5 Scope and Limitations 14
1.6 Methodology 14
1.6.1 Quantitative Phase 14
1.6.2 Qualitative Phase 15
1.7 Integration of Findings 15
Chapter 2. Literature Review 16
2.1 Introduction 16
2.2 Foundations of SaaS Business Models and Traditional Pricing 17
2.2.1 The Evolution of SaaS Economics and Subscription Logic 17
2.2.2 Historical Context: Software Cost Estimation Models 18
2.3 The Transformative Role of Artificial Intelligence in Software 19
2.3.1 AI as a Structural Shift and Market Imperative 19
2.3.2 The Unique Cost and Value Dynamics of AI 20
2.3.3 Conceptualizing AI Capabilities and Innovation 21
2.4 Challenges in Monetizing Embedded AI Features 22
2.4.1 The Complexity of AI Cost Structures 22
2.4.2 Customer Value Perception and Predictability Anxiety 23
2.4.3 Operational and Go-to-Market Hurdles 24
2.4.4 Ethical and Regulatory Complexities 25
2.5 Emerging Monetization Strategies for AI-Powered SaaS 26
2.5.1 The Dominance of Hybrid Pricing Models 27
2.5.1.1 Archetypes and Implementation 27
2.5.2 Consumption-Based Pricing and Metrics 28
2.5.2.1 Token-Based Pricing 28
2.5.2.2 Output Units and Activity Metrics 29
2.5.2.3 Prepaid Credits and Committed Spend 29
2.5.3 Outcome-Based Pricing (OBP): The Strategic Ideal 29
2.5.3.1 Structure and Examples 30
2.5.3.2 Execution Challenges and Customer Resistance 30
2.5.4 Operationalizing Monetization: Billing, Agility, and GTM 31
2.5.4.1 Re-architecting Billing and Infrastructure 31
2.5.4.2 Pricing Agility and Market Responsiveness 31
2.5.4.3 Go-to-Market (GTM) Transformation 32
2.6 Identified Deficiencies in Existing Literature, Study Contribution, and Conclusion 32
2.6.1 Identified Deficiencies in Existing Literature 32
2.6.2 Positioning the Proposed Study and Contribution 34
2.6.3 Conclusion 35
Chapter 3: Research Methodology 36

3.1 Introduction....36
3.2 Research Philosophy and Approach....36
3.3 Research Design....37
3.4 Sampling Strategy....38
3.4.1 Quantitative Sampling Strategy....38
3.4.2 Qualitative Sampling Strategy....39
3.5 Data Collection Methods....40
3.5.1 Quantitative Data Collection: Pricing Page Analysis....40
3.5.2 Qualitative Data Collection: Semi-Structured Interviews....41
3.6 Data Analysis Techniques....42
3.6.1 Quantitative Data Analysis....42
3.6.2 Qualitative Data Analysis....42
3.6.3 Integration of Quantitative and Qualitative Data....43
3.7 Validity and Reliability / Trustworthiness....43
3.7.1 Quantitative Phase: Validity and Reliability....43
3.7.2 Qualitative Phase: Trustworthiness....44
3.8 Ethical Considerations....44
3.9 Limitations of the Methodology....45
3.10 Chapter Summary....45
Chapter 4. Research Results....47
4.1 Introduction....47
4.2 Quantitative Results....48
4.2.1 Sample Characteristics....48
4.2.2 Prevalence of AI Monetisation Models....48
4.2.3 AI Packaging Strategies....49
4.2.4 Revenue-Based Comparisons....50
4.2.5 Regional Comparisons....53
4.3 Qualitative Results....53
4.3.1 Overview of Interview Sample....54
4.3.2 Emergent Qualitative Themes....54
Theme 1: Managing AI Cost Uncertainty....54
Theme 2: Customer Resistance to Pure Usage-Based Pricing....55
Theme 3: Competitive Pressure and Feature Parity....56
4.3.3 Summary of Qualitative Patterns....57
4.4 Integration of Quantitative and Qualitative Results....57
4.5 Summary....58
Chapter 5 Discussion....60
5.1 Introduction....60
5.2 Summary of Key Findings....61
5.2.1 Classification of AI Monetization Models....61
5.2.2 Prevalence by Revenue Segment....62
5.2.3 Geographic Differences....62
5.2.4 Executive Decision Drivers....63
5.2.5 Cost–Value Alignment Strategies....63
5.3 Interpretation of Findings....64
5.3.1 AI as a Structural Break in SaaS Economics....64
5.3.2 Why Hybrid Models Dominate....65
5.3.3 Psychological Anchoring and Customer Perception....65

5.3.4 Strategic Trade-offs Identified....66
5.3.5 Revenue Segment Maturity Effects....68
5.3.6 Geographic Interpretation....69
5.3.7 Evolutionary Pathways....69
5.4 Implications....71
5.4.1 Theoretical Implications....71
5.4.2 Managerial Implications....72
5.5 Limitations of the Study....74
5.5.1 Sample Limitations....74
5.5.2 Geographic Scope Constraints....75
5.5.3 Reliance on Self-Reported Performance Measures....75
5.5.4 Cross-Sectional Design....75
5.5.5 Classification Subjectivity....76
5.6 Recommendations for Future Research....76
5.6.1 Longitudinal AI Pricing Evolution Studies....76
5.6.2 Large-Scale Quantitative Performance Linkage....76
5.6.3 Experimental Pricing Studies....77
5.6.4 Enterprise vs. SMB AI Pricing Dynamics....77
5.6.5 Cost Engineering and Pricing Integration....77
5.6.6 Cross-Regional Deep Comparative Studies....77
5.7 Conclusion....78
Chapter 6. Conclusion....80
6.1 Restatement of the Research Problem and Objectives....80
6.2 Summary of Main Findings....81
6.2.1 Quantitative Findings....81
6.2.2 Qualitative Findings....82
6.3 Implications of the Research....84
6.3.1 Theoretical Implications....84
6.3.2 Managerial Implications....84
6.4 Limitations of the Study....86
6.5 Recommendations for Future Research....87
6.6 Final Thoughts and Reflections....88
Definitions of Key Terms....90
Bibliography....91
Appendix A: Quantitative Data....93
Company Distribution By Revenue....93
Pricing Strategy Distribution....94
Pricing Strategy by Revenue Range....95
Features Provided by Reviewed Companies....96
Appendix B: Qualitative Assessment....98
Questionaire....98

Table of Figures

Figure 1: Challenges in Monetizing Embedded AI Features....26
Figure 2: Pricing Strategy by Revenue....51
Figure 3: Pricing Strategies....52
Figure 4: Strategic Trade-offs of AI Pricing....68
Figure 5: 5-stage evolutionary pathway for AI monetization....71
Figure 6: 5-Step Strategic Framework....73
Figure 7: Company Distribution By Revenue....93
Figure 8: Pricing Strategy Distribution....94
Figure 9: Pricing Strategy by Revenue Range....95
Figure 10: Features Provided by Reviewed Companies....97

Index of Tables

Table 1: Inclusion and Exclusion Criteria....39
Table 2: Revenue-Based Monetization Models....51
Table 3: Pricing Strategies....52
Table 4: Company Distribution By Revenue....93
Table 5: Pricing Strategy Distribution....94
Table 6: Pricing Strategy by Revenue Range....95
Table 7: Features Provided by Companies....96

Chapter 1. Introduction

1.1 Background

The global Software as a Service market is undergoing a profound transformation driven by the rapid integration of artificial intelligence into commercial software offerings. Historically, SaaS business models evolved from perpetual licensing to subscription-based delivery, enabled by cloud infrastructure and characterised by high fixed development costs and near-zero marginal distribution expenses. This structure supported predictable recurring revenue, tiered pricing, and value metrics such as number of seats, storage limits, and feature access. Over the past decade, SaaS pricing strategies diversified further, progressing from simple subscription tiers to usage-based models and hybrid structures designed to align more closely with customer value and product usage patterns. However, these developments occurred largely in a pre-generative AI environment and were not designed to accommodate the new economic dynamics introduced by AI-driven capabilities.

The infusion of AI represents a structural shift rather than a feature enhancement. AI transforms software from a passive tool into an active collaborator capable of generating insights, content, predictions, and operational decisions. Generative AI, in particular, has accelerated this shift, with industry analysts describing it as one of the most disruptive forces in enterprise software. Valued at over USD 266 billion in 2024 and projected to exceed USD 1.1 trillion by 2032, the industry's growth is increasingly intertwined with the AI market, which is itself expanding at a compound annual growth rate (CAGR) of 31.5% (Fortune Business Insights, 2025). This convergence signals that AI will not merely supplement SaaS offerings but will redefine expectations for functionality, customer value, and commercial models.

A parallel trend is visible in emerging markets such as India, where SaaS revenues are projected to hit $50 billion ARR by 2030 (Bessemer Venture Partners, 2023). At the same time, generative artificial intelligence (GenAI) has emerged as “the most disruptive force in enterprise software since the dawn of SaaS” (Khanna et al., 2025), driving a rush of new AI-enabled features into SaaS products. Leading analysts observe that AI’s value-driven capabilities are “driving innovation in outcome-based, consumption-based and hybrid strategies” and forcing a rethink of traditional pricing models (L.E.K. Consulting, 2025).Together, these global and regional trends highlight the increasing strategic importance of AI within the SaaS sector.

Yet despite the enthusiasm surrounding AI, SaaS companies face substantial challenges in monetizing AI capabilities effectively. Most traditional SaaS pricing approaches presume negligible marginal costs. In contrast, AI-powered features introduce variable, usage-based expenses that scale with user activity, particularly for large language model inference, which cloud providers price on a per-token or per-inference basis. This change disrupts long-standing assumptions about SaaS unit economics, gross margins, and cost-to-serve. Whereas increased usage previously signalled product success and improved profitability, high AI inference volumes can now result in margin compression or financial losses if pricing is misaligned with actual costs.

In addition to economic challenges, customers' value perceptions further complicate AI monetization. While AI promises significant productivity gains and business impact, many customers struggle to quantify this value or compare it meaningfully against costs. Their reference points are often shaped by low-cost consumer AI applications, which creates an anchoring effect that depresses willingness to pay for enterprise-grade AI features. Meanwhile, competitive pressures push SaaS vendors to include AI functionality to remain relevant, even when monetization strategies are unclear or underdeveloped.

These dynamics create uncertainty for SaaS companies attempting to design sustainable pricing and packaging strategies. Industry analyses from consulting firms and technology research organisations identify emerging models such as usage-based metering, token-based pricing, AI add-on subscriptions, consumption credits, and outcome-based structures. However, these insights remain largely descriptive, fragmented, and lacking empirical grounding. Existing academic literature on SaaS pricing predates these developments and focuses on traditional value metrics that do not reflect AI-driven variability. This has left practitioners without reliable benchmarks or validated frameworks for designing AI pricing structures.

Against this backdrop, there is a clear need for systematic empirical research to examine how SaaS firms are approaching AI monetization, which models are gaining adoption, and what strategic considerations shape executive decision-making. Understanding these dynamics is essential for developing pricing strategies that align customer value, competitive positioning, and the new cost structures introduced by AI.

1.2 Problem Statement

The central problem confronting the Software as a Service (SaaS) industry is the structural misalignment between the established economic logic of SaaS business models and the new cost and

value dynamics introduced by embedded Artificial Intelligence (AI). SaaS companies are struggling to formulate and implement pricing and packaging strategies for AI features that are simultaneously profitable, scalable, and aligned with customer value perceptions. This challenge stems from a confluence of technological, economic, and psychological factors that render traditional pricing frameworks inadequate.

The background of this problem lies in the fundamental architectural shift of software. The traditional SaaS model, characterized by high fixed costs and near-zero marginal costs of distribution, optimized for predictable recurring revenue through per-seat subscriptions (BCG, 2025). In contrast, AI-powered SaaS introduces a significant and variable marginal cost. The computational expense of running LLMs, often priced on a per-token basis by providers like OpenAI, AWS, and Azure, means that every user query or AI-driven action incurs a direct, tangible cost that scales linearly with usage (Valueships, 2025). This reality forces SaaS vendors into a precarious position: they must now price for consumption to cover these costs, a model that can conflict with customer preferences for budget predictability (McKinsey & Company, 2025).

The potential causes of this problem are multifaceted. Firstly, the **cost structure of AI is inherently complex and dynamic**. Beyond the direct costs of API calls (per-token), there are substantial expenses related to hosting, fine-tuning models, data management, and the specialized talent required for development and maintenance (SaaStr / Unusual Ventures, 2025). Secondly, **customer value perception is difficult to anchor**. While AI features can unlock immense productivity gains, estimated to add up to $4.4 trillion in annual global economic value, communicating this return on investment (ROI) is a significant challenge (McKinsey & Company, 2025). Many customers struggle to quantify the value of "smarter" software, and their price sensitivity is often anchored by low-cost consumer AI applications (e.g., ChatGPT Plus at ~$20/month), creating a disconnect with the higher costs of enterprise-grade solutions. Thirdly, there is a **lack of established best practices**. The technology is evolving so rapidly that the industry is in a state of widespread experimentation with various models; including usage-based, outcome-based, credit-based, and complex hybrid structures; without a clear consensus on what works (BCG, 2025).

1.3 Objectives

This study pursues a set of primary and secondary objectives that collectively aim to produce a comprehensive and practical understanding of AI monetization strategies within the B2B SaaS sector.

1.3.1 Primary Objective

The primary objective of this study is to conduct a mixed-methods empirical investigation into the pricing and packaging strategies for Artificial Intelligence (AI) features adopted by B2B Software as a Service (SaaS) companies with annual revenues between $1 million and $100 million. This research aims to systematically map and classify the prevalent monetization models (e.g., subscription add-ons, usage-based, credit-based, outcome-based, and hybrid structures). Furthermore, through qualitative inquiry with senior executives, the study seeks to uncover the underlying strategic, behavioral, and competitive rationales that inform these choices, culminating in the development of a practical framework to aid executive decision-making on AI monetization.

1.3.2 Secondary Objectives

- To identify and classify the primary pricing and packaging strategies SaaS companies are using to monetize embedded AI features.
- To determine the prevalence of different AI monetization models (e.g., seat-based, usage-based, token-based, outcome-based, hybrid) across the SaaS industry within the specified revenue segment.
- To analyze differences in AI monetization approaches across SaaS company sub-segments based on annual revenue ($1-10MM; $10-50MM; $50-100MM).
- To explore the strategic considerations, perceived challenges, and competitive factors that influence executives' choices in designing AI pricing and packaging.
- To investigate how SaaS companies attempt to balance variable AI costs with customer demand for price predictability and value alignment.
- To propose a practical, evidence-based framework that can support executive decision-making on AI pricing strategy.

1.4 Significance

The rapid infusion of AI into SaaS products represents one of the most significant technological and business model shifts in the last decade. However, the strategies for monetizing this shift remain nascent, poorly understood, and largely undocumented in a systematic, academic manner. This study is significant because it addresses this critical knowledge gap with a rigorous, empirically grounded

approach, offering substantial contributions to theory, practice, and industry discourse.

This study represents an improved effort to solve the problem of AI monetization for several reasons. Firstly, it moves beyond the anecdotal evidence and opinion-based articles that currently dominate the conversation. By employing a **mixed-methods research design**, it combines the breadth of a large-scale quantitative analysis of pricing pages with the depth of qualitative executive interviews. This approach provides a more holistic and validated understanding than either method could achieve alone, capturing both the "what" and the "why" of current strategies. Secondly, its focus on a specific, highly relevant segment of the market, SaaS companies with $1M to $100M in annual revenue ensures the findings are targeted and actionable. This segment is mature enough to have established pricing strategies but agile enough to be actively experimenting, making it an ideal crucible for observing emerging best practices.

This research is significant for both scholarship and practice. From an **academic** perspective, it bridges the gap between technology adoption and business model innovation. The rise of AI in software constitutes a structural change that challenges classic pricing theory. By studying how firms align **value creation** with pricing in an AI context, the project will contribute new insights to the literature on digital goods, platform economics, and value-based pricing. It will also enrich understanding of cross-border variations (US vs India) in tech strategy, a relatively unexplored area in SaaS research. Finally, by combining quantitative and qualitative data, the study can inform and extend theoretical understanding (e.g. on usage-based pricing, outcome metrics) in a real-world setting.

From a **practical** standpoint, the findings will guide SaaS entrepreneurs, product managers and pricing strategists. Pricing is one of the most critical yet fraught decisions in SaaS (Bain & Company, 2022; McKinsey & Company, 2025). SaaS companies worldwide are rapidly adding AI features but lack benchmarks or best practices on monetization. This research will systematically reveal which pricing/packaging models are most common, and via interviews illuminate what drives those choices. For example, executives may share how they negotiated trade-offs between revenue growth and gross margins, or how competitors influenced their decisions. Such knowledge can help startups and established firms avoid pitfalls (e.g. margin traps or customer churn) when deploying AI. It will also aid investors and consultants who advise SaaS firms in setting pricing strategy.

Economically, effective AI monetization is vital for sustaining SaaS innovation. Given that AI development is expensive and often funded by venture capital (especially in India’s booming SaaS ecosystem (Bessemer Venture Partners, 2023)), companies need to demonstrate revenue models that

justify those investments. Poor pricing can undermine growth or lead to unsustainable burn. Thus, by identifying viable pricing frameworks, this study supports the broader success of the SaaS industry. In short, the study's findings will provide evidence-based guidance on a problem that industry leaders currently regard as a **"board-level imperative"** (Metronome, 2025), making it a highly valuable contribution to both management practice and policy.

1.5 Scope and Limitations

The scope of this study is deliberately focused on B2B SaaS companies with annual recurring revenue between USD 1 million and USD 100 million. This segment excludes early-stage startups that have not yet developed stable pricing models and large enterprises whose pricing structures often depend on custom-negotiated contracts that lack transparency. The study also focuses specifically on companies that offer identifiable AI features, thereby ensuring the relevance of findings to AI monetization.

Finally, the study focuses on monetization of embedded AI features rather than broader organisational transformation or AI-enabled product development processes. It does not evaluate the technical performance of AI systems but rather the commercial strategies used to price them.

1.6 Methodology

This study will employ a mixed-methods sampling design, utilizing distinct strategies for the quantitative and qualitative phases to ensure comprehensive data collection.

The population for this research consists of Software as a Service (SaaS) companies globally. The sampling frame is further defined as companies with annual recurring revenue (ARR) between $1 million and $100 million USD. This revenue range is selected to exclude early-stage startups still in price discovery and large enterprises whose privately negotiated contracts would obscure standard pricing models.

1.6.1 Quantitative Phase

Quantitative (Pricing Page Analysis): We will systematically collect pricing and packaging data from the websites of the sampled companies. This will involve manually or automatically scraping pricing pages, download links, or product FAQs. Key variables to record include: pricing model (subscription, usage-based, tiered, etc.); value metric (per user, per token, per outcome, etc.); number of tiers/plans; presence of AI features and how they are labeled; pricing levels (list prices); and any usage caps or overage rules. We will code these data into categories (e.g. "Usage-based (AI tokens)", "Seat + AI add-

on", "Outcome-based") based on a predefined codebook. Descriptive statistics will then be used to identify patterns, for example, the percentage of companies using hybrid pricing versus pure subscription.

1.6.2 Qualitative Phase

Qualitative (Executive Interviews): Using a semi-structured interview guide, we will interview executives via video or phone (each ~45–60 minutes). Topics will cover how the company decided on an AI pricing strategy, which factors (e.g. cost, competition, customer feedback) influenced it, and what challenges were faced. We will ask specifically about packaging decisions: for instance, why AI was bundled or sold separately, how usage is metered, and how prices were set. Interviews will be recorded (with consent), transcribed, and analyzed using thematic coding.

1.7 Integration of Findings

The mixed-methods approach ensures that quantitative insights into pricing prevalence are complemented by qualitative exploration of strategic motivation. By integrating these findings, the study will produce a holistic framework that helps SaaS leaders evaluate monetization options aligned with their organisational context.

Chapter 2. Literature Review

2.1 Introduction

The pervasive integration of Artificial Intelligence (AI) and, specifically, Generative AI (GenAI), into commercial software offerings marks a **structural shift** for the Software as a Service (SaaS) industry, fundamentally challenging the established economic logic of traditional SaaS models. The magnitude of this transformation is significant, with the global SaaS market, valued at over **USD 266 billion in 2024** and projected to exceed **USD 1.1 trillion by 2032**, having its growth increasingly intertwined with the expanding AI market. This disruption is often viewed by industry leaders as transformative, signaling that AI will redefine expectations for functionality and commercial models.

The convergence of AI and SaaS has created a central problem: a structural misalignment between traditional predictable SaaS pricing approaches (optimized for predictable recurring revenue through per-seat subscriptions) and the new, variable cost dynamics introduced by embedded AI features. Unlike traditional SaaS, where adding users had negligible marginal costs, AI-powered features introduce significant and variable marginal costs that scale directly with user activity (such as per-token pricing for Large Language Model inference). This reality forces companies to urgently seek monetization strategies that are both profitable and scalable, as high usage can now lead to margin erosion or financial losses if pricing is misaligned.

This challenge is complicated by the inherently complex and dynamic nature of AI cost structures, extending beyond API calls to include expenses for hosting, fine-tuning models, and acquiring specialized talent. Furthermore, customer price sensitivity is often anchored by low-cost consumer AI applications (e.g., ChatGPT Plus), and enterprise buyers exhibit anxiety over unpredictable usage costs, complicating efforts to monetize effectively.

While fragmented industry reports document emerging hybrid and consumption-based monetization models, the existing literature is largely descriptive and lacks the methodological rigor of systematic, empirical research needed to provide reliable benchmarks and validated decision frameworks. This comprehensive literature review is thus essential to synthesize the current state of SaaS pricing theory and AI cost dynamics, define the spectrum of emerging monetization strategies, and ultimately, justify the need for a targeted empirical investigation into mid-sized B2B SaaS firms (US$1M to $100M ARR), thereby contributing critical insights to academic and practical understanding.

2.2 Foundations of SaaS Business Models and Traditional Pricing

The economic architecture of Software as a Service (SaaS), prior to the widespread integration of Artificial Intelligence (AI), was established during a transition from perpetual licensing to a subscription-based delivery model facilitated by cloud computing. This shift redefined how software revenue was generated and how costs were understood, laying a foundation that the current AI revolution fundamentally challenges.

2.2.1 The Evolution of SaaS Economics and Subscription Logic

The traditional SaaS model operated under an established economic logic characterized by high fixed development costs and near-zero marginal costs of distribution. Once the software was built, distributing it to an additional user incurred negligible expense, enabling profitable scaling. The primary goal of this model was the attainment of predictable recurring revenue (ARR) through stable contracts, pricing tiers, and defined value metrics .

Pricing strategy in this traditional B2B SaaS environment was intrinsically linked to maximizing value capture from customers over time(DiVA Portal, 2015). The underlying philosophy often centered on value-based pricing (VBP), which dictates that prices should be set primarily based on the perceived or estimated value of a product or service to the customer, rather than strictly on production costs or competitor rates(DiVA Portal, 2015; Maxio, n.d.),. Crucially, customer price perception extended beyond mere cost; customers frequently interpreted the price as an indicator of service quality. Accordingly, successful monetization depended on clearly quantifying and communicating the Return on Investment (ROI) derived by the customer (DiVA Portal, 2015; Maxio, n.d.).
The models developed to monetize this low-marginal-cost paradigm were predominantly fixed and predictable:

1. Subscription/Per-Seat Models: These were the dominant models, utilizing per-user or per-seat pricing, which provided simplicity for customers accustomed to predictable procurement workflows and offered vendors predictable revenue streams. This metric aligned value capture with employee usage, reflecting the software's ability to enhance human productivity (Bain & Company, 2025).

2. Tiered Models: These models offered customers a range of predetermined bundles of products or features, often organized according to ascending prices and complexity, to suit different customer sizes and needs. The tiered approach provided predictability in revenues for the vendor compared to pure consumption models (DiVA Portal, 2015).

3. Freemium and Trial Models: These strategies, often time- or feature-based, aimed to lower the initial commitment barrier, allowing potential customers to test the product and realize its value before committing to a paid subscription(DiVA Portal, 2015).

Before AI, the market had already begun shifting toward hybrid models, integrating aspects of usage metrics into the subscription core, as metrics like seats were becoming increasingly misaligned with the complexity and variety of digital products,. However, this evolution occurred largely outside the dynamic cost pressures now introduced by AI features (BCG, 2025; Zuora, 2025).

2.2.2 Historical Context: Software Cost Estimation Models

The challenge of determining pricing in the pre-AI SaaS world largely focused on accurately estimating the extensive fixed, upfront costs associated with developing the software. This area of research relied on algorithmic cost estimation models designed to predict the effort (man-months or hours) required for software development projects (Kemerer, 1986).

A major study empirically validated four of the most popular algorithmic models used to estimate software costs: SLIM, COCOMO, Function Points, and ESTIMACS. This historical research categorized these models primarily based on their key input metrics, determining whether they relied on Source Lines of Code (SLOC) or non-SLOC inputs (Kemerer, 1986).

1. SLOC-Based Models (SLIM and COCOMO): These models, such as the Constructive Cost Model (COCOMO) developed by Barry Boehm, utilize delivered source instructions (DSI) or KSLOC (SLOC stated in thousands) as their primary input to predict effort. SLOC was an early metric selected by researchers due to its apparent quantifiability. However, the reliability of these models hinged heavily on the accuracy of the SLOC estimates, which are difficult to obtain ex ante (before a project begins) (Kemerer, 1986).

2. Non-SLOC Models (Function Points and ESTIMACS): These methods sought inputs easier to estimate earlier in the project life cycle. The Function Points method, validated on independent data, captures functionality at a more macro level than SLOC, focusing on factors like the number of input transaction types and unique reports. This approach avoids complexities introduced by implementation differences, such as programming language(Kemerer, 1986).

The empirical validation of these models demonstrated significant challenges when applying them outside their original development environments. Uncalibrated models exhibited massive average error rates, ranging from 85% to 772%, with many falling in the 500-600% range. While the models showed improved predictive power after careful local calibration (with the best explaining up to 88% of actual

effort variation), the study concluded that, as models of the overall software development process, they failed to sufficiently reflect the underlying factors affecting productivity. This entire historical context centered entirely on managing and estimating the fixed investment costs (man-months of effort) and is fundamentally insufficient for tackling the dynamics of AI, which introduces variable and transactional costs that scale linearly with usage (Kemerer, 1986).

2.3 The Transformative Role of Artificial Intelligence in Software

The integration of Artificial Intelligence (AI) and, specifically, Generative AI (GenAI), into commercial software represents a profound **structural shift** that fundamentally redefines the relationship between software vendors and their customers. This technological evolution moves software beyond its historical function as a passive tool that primarily enhances human productivity, transforming it into an **active collaborator** capable of generating insights, content, predictions, and operational decisions. This shift challenges the core economic premises of traditional SaaS established in the pre-AI era, where business models were optimized around high fixed development costs and near-zero marginal distribution costs (McKinsey & Company, 2025).

2.3.1 AI as a Structural Shift and Market Imperative

The infusion of AI is widely regarded by industry analysts not merely as a feature enhancement but as **"the most disruptive force in enterprise software since the dawn of SaaS"**. This convergence of technologies has made effective AI monetization a **C-level imperative** for SaaS firms seeking to remain competitive and unlock new growth potential (Khanna et al., 2025).

The market opportunity driven by this shift is enormous, rapidly expanding the total addressable market for software offerings. The global Software as a Service market was valued at over **USD 266 billion in 2024** and is projected to exceed **USD 1.1 trillion by 2032**, exhibiting a compound annual growth rate (CAGR) of 20.00% (Fortune Business Insights, 2025). This growth is deeply intertwined with the broader AI market, which is expanding at an even faster CAGR of 31.5% (Fortune Business Insights, 2025). Moreover, acceleration of AI adoption is evident, with 46% of surveyed companies capturing financial impact or seeing productivity benefits at scale from AI in 2025, up from 33% just one year prior (Khanna et al., 2025).

Regional markets reflect this urgency. India's Software as a Service ecosystem, for instance, is projected to reach **$50 billion ARR by 2030** (Bessemer Venture Partners, 2023). Indian software companies possess an **efficiency advantage** and are positioned to expand into global leadership by building additional products faster and targeting market segments effectively. This structural

transformation signals that AI features will not merely supplement existing SaaS but will redefine customer expectations for functionality and commercial models (Bessemer Venture Partners, 2023).

2.3.2 The Unique Cost and Value Dynamics of AI

The fundamental economic challenge posed by embedded AI lies in its disruption of the traditional SaaS unit economics model that assumed negligible marginal costs for adding users (McKinsey & Company, 2025; Valueships, 2025). In stark contrast, AI features introduce **significant and variable marginal costs** that scale linearly with user activity. The computational expense of running Large Language Models (LLMs), particularly for inference (when the model is used), is frequently priced by cloud providers (such as OpenAI, AWS, and Azure) on a **per-token or per-inference basis** (McKinsey & Company, 2025; Valueships, 2025).

This reality forces SaaS vendors into a precarious financial position: product success, characterized by high user engagement and usage volume, can paradoxically lead to **margin compression or financial losses** if the pricing strategy is misaligned with the variable cost structure (Valueships, 2025). The financial exposure is considerable, given that GenAI gross profit margins are estimated to be, on average, 10 to 30 percentage points lower than those of traditional SaaS offerings. Operational costs can be astronomical; for example, OpenAI reportedly spends an estimated **$700,000 daily to run ChatGPT**, demonstrating that even highly successful market players face challenges where operational costs threaten to exceed revenue growth (Zuora, 2025).

The total cost structure of AI is highly dynamic and multi-layered, encompassing both substantial fixed and scaling variable expenses (BCG, 2025):

- **Variable Costs:** These include direct API call costs (per-token or per-inference), which scale with user consumption (SaaStr / Unusual Ventures, 2025).
- **Fixed/Overhead Costs:** These are significant upfront investments in R&D, such as model training, pre-training, re-training, and fine-tuning specialized models. The cost of training models requires immense computational resources; Meta's LLaMA 2, for instance, required over **3 million GPU hours**, costing approximately $4 million solely in hardware usage. Additional overhead includes costs associated with curating high-quality data and compensating **specialized talent** (Coherent Solutions, 2025).

Despite these costs, the **potential opportunity is massive**, as previous McKinsey research estimates that AI features have the potential to add **up to $4.4 trillion in annual global economic value** (Khanna et al., 2025). To capture this value, AI's capabilities are **"driving innovation in outcome-**

based, consumption-based and hybrid strategies," necessitating a definitive shift away from traditional fixed-fee subscriptions (L.E.K. Consulting, 2025).

2.3.3 Conceptualizing AI Capabilities and Innovation

To systematically manage and monetize this transformative technology, it is essential to understand the underlying structure of AI systems and their impact on existing business competencies.

AI System Building Blocks: AI systems are essentially computational agents designed to act intelligently. They can be systematically broken down into six core **building blocks** (Paschen et al., 2020):

1. **Structured Data:** Organized information foundational to business analytics (e.g., inventory figures, sales data).
2. **Unstructured Data:** Non-standardized data (e.g., blog posts, audio, video) that AI is uniquely equipped to handle.
3. **Preprocesses:** Functions like data cleaning, Natural Language Understanding (NLU), and computer vision, necessary to make data usable by the AI system.
4. **Main Processes:** The execution layer involving intelligent behaviors such as problem solving, reasoning (deductive/inductive), and machine learning/deep learning.
5. **Knowledge Base:** The storage layer for past data, information, and implicitly derived knowledge used to inform subsequent intelligent behavior.
6. **Value-Added Information Outputs:** The results, conveyed as meaningful information (e.g., natural language generation, image generation, or actions via robotics).

Innovation Typology: The way AI affects the market and organizational structure can be analyzed using a typology of AI-enabled innovation, classified along two major dimensions (Paschen et al., 2020):

1. **Innovation Boundaries:** Distinguishes between **product-facing innovations** (influencing a firm's offerings, e.g., AI in navigation apps) and **process-facing innovations** (influencing internal operations, e.g., AI automating media buying).
2. **Organizational Competencies:** Describes whether the innovation is **competence-enhancing** (amplifying current knowledge and skills) or **competence-destroying** (rendering existing skills obsolete).

The current trend of embedding AI features into SaaS platforms typically constitutes a **competence-enhancing product innovation**, improving performance while building upon existing competencies.

For instance, AI assisting a biochemist enhances their ability to crunch data, just as AI in a navigation app improves the driving experience without obsoleting the core driving skill. However, the shift towards AI agents and task automation suggests an increasing move toward competence-destroying innovations that require SaaS vendors to radically adapt their fundamental business models and competencies to capture value (Khanna et al., 2025).

2.4 Challenges in Monetizing Embedded AI Features

The convergence of Artificial Intelligence (AI) and Software as a Service (SaaS) presents a **structural misalignment** between the predictable economic models of legacy software and the new dynamics of intelligent, constantly evolving digital features. Monetizing embedded AI features effectively requires navigating a complex minefield of financial exposure, customer psychology, operational inertia, and ethical risks. These challenges render traditional fixed-fee pricing inadequate and necessitate a wholesale rethink of monetization strategy (Bain & Company, 2025).

2.4.1 The Complexity of AI Cost Structures

The most immediate and disruptive challenge stems from AI's alteration of the fundamental unit economics of SaaS. The traditional model was predicated on high fixed development costs offset by **near-zero marginal costs of distribution** (BCG, 2025). In contrast, AI features introduce **significant and variable marginal costs** that scale directly with user activity (McKinsey & Company, 2025).

The computational expense of running large language models (LLM inference) is often priced by cloud providers (such as OpenAI, AWS, and Azure) on a **per-token or per-inference basis**,. This structural reality means that every user query incurs a direct, tangible cost that scales linearly with usage. This contrasts sharply with the low marginal cost of traditional SaaS models, where distributing to an additional user incurred negligible expense (Zuora, 2025).

This dynamic creates a significant paradox: while increased usage previously signaled success and led to higher profitability, high AI inference volumes can now result in **margin compression or financial losses** if the pricing is misaligned with the consumption costs,. GenAI gross profit margins are estimated to be, on average, **10 to 30 percentage points lower** than those for traditional SaaS offerings. This cost reality is evident in industry examples, such as OpenAI reportedly spending an estimated **$700,000 daily to run ChatGPT**, demonstrating that even market leaders face immense operational costs that can exceed revenue growth, despite generating revenues potentially surpassing $2 billion annually (Valueships, 2025; Zuora, 2025) .

The total cost of developing and sustaining AI goes far beyond the direct variable costs of inference:

1. **Complexity of the AI Model:** This can account for **30–40%** of the total project cost. Building or training large-scale models requires **vast amounts of data** and significant computing power, often necessitating specialized hardware like GPUs or TPUs,. For instance, training Meta's LLaMA 2 models required over **3 million GPU hours**, incurring hardware usage costs estimated at approximately **$4 million** (Coherent Solutions, 2025).
2. **Data Requirements:** Data collection and preparation account for **15–25%** of the total cost,. Most businesses start without sufficient training data. Generating the required 100,000 data samples for a complex machine learning project can cost around **$70,000**, and cleaning data to remove errors and biases can take between **80 and 160 hours** for that same size dataset (Coherent Solutions, 2025).
3. **Infrastructure and Technology:** Representing **15–20%** of total AI development costs, this includes expenses for cloud computing resources, hosting, data storage, and the specific architecture needed for handling real-time data processing (Coherent Solutions, 2025).
4. **Fixed Overhead and Talent:** The costs also include substantial R&D expenses related to **model (re)training, fine-tuning models**, and highly competitive compensation for **specialized talent**,,. Even a small AI development team can cost upward of **$400,000 per year** in technology development costs alone. Average salaries for data scientists and ML engineers range from **$120,000 to over $220,000 per year** in the US (Coherent Solutions, 2025).

This multi-layered, dynamic, and substantial cost structure mandates that SaaS companies must price for consumption to cover their financial exposure, a requirement that directly clashes with deep-seated customer preferences.

2.4.2 Customer Value Perception and Predictability Anxiety

The monetization challenge is intensely amplified by psychological and market dynamics that complicate value capture, even when AI features promise substantial productivity gains (McKinsey & Company, 2025).

Firstly, despite AI's potential to add **up to $4.4 trillion in annual global economic value**, communicating and quantifying this Return on Investment (ROI) to customers remains a significant hurdle,,. Few companies successfully translate this value: only approximately **30%** of firms have published quantifiable ROI figures in dollar terms from real customer deployments. Customers frequently **struggle to quantify the value** of "smarter" software (SaaStr / Unusual Ventures, 2025).

Secondly, customer expectations for AI pricing are heavily influenced by the **anchoring effect** established by ubiquitous consumer applications, such as ChatGPT Plus, which often cost around

$20/month,. This low price point creates a significant disconnect when vendors attempt to charge higher rates appropriate for complex, enterprise-grade AI features and their underlying costs (Khanna et al., 2025).

Thirdly, and most critically, customer **anxiety over unpredictable costs** stands as the single biggest barrier to enterprise adoption,. Enterprise buyers lack confidence in calculating their total financial exposure and fear receiving "surprise bills" from metered usage,,. This anxiety is so profound that customers often **actively avoid using AI features** even when vendors offer free credits, out of fear of getting locked into an unpredictable consumption model (albertsaro, n.d.).

This pervasive desire for budget stability means enterprise buyers frequently **push back on AI line items unless they are flat-rate or capped**,. For SaaS vendors, **predictability, not the lowest price point, drives enterprise adoption**,. To overcome this resistance, companies must provide clear expectations through mechanisms like caps, rollovers, or flat rates to unlock usage. The sales process is further complicated because vendors often ask customers to commit to spending **before** they have seen commensurate savings or realized the ROI, creating a necessary financial gap that must be bridged with incentives and flexible contracts (Bain & Company, 2025).

2.4.3 Operational and Go-to-Market Hurdles

Transitioning from stable per-seat pricing to dynamic consumption or hybrid models exposes deep-seated technical and organizational shortcomings within established SaaS firms.

1. **Billing and Infrastructure Fragmentation:** Most incumbent SaaS companies lack the necessary **product telemetry** (real-time usage tracking) and the robust IT, billing, and finance infrastructure required to handle dynamic usage data at scale,,. Pricing flexibility is fundamentally constrained by billing capability. Large companies often operate with fragmented systems, running separate stacks for Product-Led Growth (PLG) (e.g., Stripe) and Sales-Led Growth (SLG) (e.g., CPQ, ERP), neither of which integrates usage pricing cleanly,. This often results in pricing logic being inefficiently embedded in **product code**, making it non-auditable and difficult to verify, a challenge highlighted by the difficulty of internal finance teams in explaining how calculated revenue figures were reached (Bain & Company, 2025).
2. **Sales and Enablement Gap:** The Go-to-Market (GTM) team faces a steep learning curve. Sales representatives are traditionally trained and incentivized to sell predictable **seats** or subscription contracts,,. The introduction of complex concepts like metered billing, outcome-based metrics, and overage policies requires entirely new capabilities, playbooks, and **ROI calculators**,,. Sales reps often revert to selling **flat rates by default** because they lack the

ability to confidently forecast variable consumption, thus undermining the core strategy of capturing consumption value (Bain & Company, 2025).

3. **Cross-Functional Alignment and Change Management:** Executing usage-based AI pricing requires tight coordination across the entire organization (Product, Finance, and GTM),,. This shift necessitates rewriting sales compensation KPIs to reward consumption and expansion rather than just upfront deal size,. Furthermore, successfully transitioning to AI-enabled operations involves significant change management: successful pilots suggest organizations should allocate **$3 to change management** (including training, performance monitoring, and engineering support) for **every $1 spent on model development**. Companies like Monte Carlo, facing "spiky" and unpredictable AI usage, demonstrated the necessity of this shift by adopting **daily revenue tracking** as a real-time, shared performance metric across teams (Bain & Company, 2025; Khanna et al., 2025; L.E.K. Consulting, 2025).

2.4.4 Ethical and Regulatory Complexities

Beyond operational hurdles, the sophistication of AI pricing systems introduces novel **ethical complexities** that challenge fundamental concepts of market fairness, consumer protection, and accountability (IJCESEN Journal, 2025).

1. **Algorithmic Discrimination and Bias:** AI pricing mechanisms are built on machine learning algorithms trained on vast historical datasets,. If this training data contains **embedded biases** reflecting past societal discrimination (e.g., historical or representation bias), the pricing algorithms can perpetuate and systematize unfair treatment of specific demographic or socioeconomic groups across millions of transactions. This risk is heightened by **geographic discrimination**, where location-based data features (often seemingly neutral) act as proxies for sensitive demographic information, creating economic barriers for consumers in certain regions (IJCESEN Journal, 2025).
2. **Transparency and the Black Box Problem:** Modern AI pricing algorithms frequently operate as a **complex "black box,"** lacking interpretability even for their creators. The complex, multi-layered structures of deep neural networks obscure the causal relationship between input data and the final pricing decision,,. This opacity creates **information asymmetry** where businesses possess detailed consumer profiles and predictive models (e.g., predicting willingness to pay) while consumers remain unaware of how their data influences the price they encounter for identical goods. This deficit makes it difficult for consumers and regulators to understand or challenge pricing decisions (IJCESEN Journal, 2025).

3. **Accountability and Regulatory Gaps:** Accountability is challenging due to the **multi-layered structure of responsibility** involved in AI development, spanning executive leadership, data scientists, business strategists, and software engineers,. Determining responsibility when discriminatory or harmful outcomes occur is difficult because the causal chain is distributed across organizational and technical domains. Furthermore, current legal and regulatory frameworks often struggle to address algorithmic discrimination,, lacking the technical specificity required to govern machine learning bias, particularly given the rapid pace of technological change and the absence of established, industry-wide best practices. This regulatory lag creates uncertainty for businesses and limits recourse for consumers who experience unfair treatment (IJCESEN Journal, 2025).

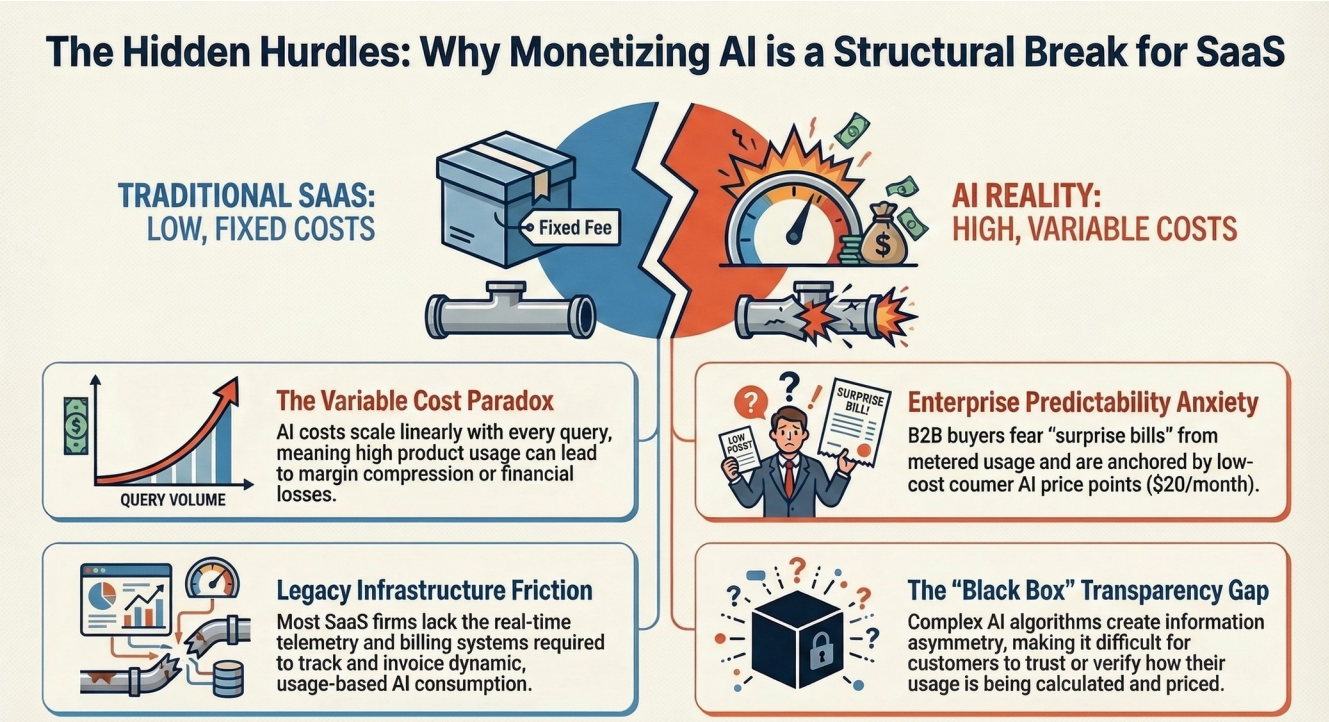


Figure 1: Challenges in Monetizing Embedded AI Features

The necessity of aligning costs with usage must thus be balanced against the inherent complexity of AI systems, pervasive customer anxiety over predictability, and the non-trivial organizational and ethical transformations required for long-term viability (IJCESEN Journal, 2025; ResearchGate, 2025).

2.5 Emerging Monetization Strategies for AI-Powered SaaS

The technological disruption caused by the embedding of Artificial Intelligence (AI) and Generative AI (GenAI) into Software as a Service (SaaS) products has rendered traditional fixed-fee pricing inadequate, compelling vendors to adopt dynamic, hybrid, and consumption-aligned models. This shift

is necessary because the structural economic assumptions of legacy SaaS - namely, near-zero marginal costs of distribution - are directly contradicted by the **significant and variable marginal costs** of AI inference, such such as per-token pricing for large language model (LLM) usage (Valueships, 2025; Zuora, 2025) . Monetization in this new era has become a **"board-level imperative"**, focused on managing financial exposure while simultaneously addressing the customer demand for price predictability (Metronome, 2025).

This section systematically reviews the emerging monetization strategies observed across the AI-SaaS ecosystem, classifying prevalent models, detailing the shift towards specialized metrics, and outlining the necessary organizational and technological foundations required for successful deployment.

2.5.1 The Dominance of Hybrid Pricing Models

The most common and effective solution adopted by SaaS companies seeking to monetize AI features is the **hybrid pricing model**, which systematically combines the traditional stability of subscription fees with variable components linked to AI usage or consumption. This blending is viewed as an ideal middle ground because it successfully navigates the core conflict between the vendor's need to cover high, variable costs and the customer's pervasive **anxiety over unpredictable spending** (Metronome, 2025; Zuora, 2025).

Hybrid models are demonstrably advantageous, with businesses utilizing them reporting higher Year-over-Year Annual Recurring Revenue (YoY ARR) growth compared to those relying solely on other models. This approach is not merely a tactical maneuver but a necessity driven by financial reality; many AI features were packaged with usage pricing not by aggressive strategy, but as a **defensive measure** to avoid immediate margin erosion due to the variable compute costs (Zuora, n.d.).

2.5.1.1 Archetypes and Implementation

The literature identifies several distinct archetypes of hybrid structures being deployed across B2B SaaS:

1. **Subscription Add-ons or Tiered Bundling:** The simplest approach involves embedding AI features directly into existing subscription tiers, often by simply increasing the per-seat price (Bain & Company, 2025; L.E.K. Consulting, 2025). This maintains familiar procurement workflows (Khanna et al., 2025). However, a more sophisticated approach bundles features into tiered plans that provide increasing levels of access and complexity (BCG, 2025; Valueships, 2025). For example, a subscription plan might come with a flat fee but unlock features corresponding to a certain **model complexity** or response time requirement.

2. **Capped or Threshold Models (Buckets):** This model offers the customer the predictability they crave by setting clear **usage caps or thresholds** for AI features on top of a stable base subscription fee. Customers pay a fixed fee for a predefined volume (a "bucket" or "allowance") of credits or tokens. Once the cap is exceeded, the customer then purchases additional "buckets" or faces overages. For instance, Box's Enterprise Plus plan includes 20 GenAI queries per user per month, complemented by a pooled company-wide allowance of 2,000 queries, with the option to purchase more (Bain & Company, 2025; Khanna et al., 2025; Zuora, 2025).
3. **Premium Copilot Pricing:** AI features designed to augment human productivity (often termed "copilot-style" offers) are frequently monetized via a substantial premium layered onto the base subscription. This premium often ranges from **30% to 110% above the base per-seat cost**. Microsoft's Copilot, for example, is priced at 60%-70% of the base product fee. This premium reflects the perceived productivity gains delivered by the AI feature, such as GitHub Copilot reporting developers coding up to 55% faster (L.E.K. Consulting, 2025).

This hybridization allows vendors to capture revenue that directly aligns with the economic value perceived by customers (value-based pricing) while ensuring cost recovery, especially for power users whose variable consumption would otherwise compromise margins in a flat-rate model (BCG, 2025; Valueships, 2025).

2.5.2 Consumption-Based Pricing and Metrics

Consumption-based pricing models involve customers paying strictly based on actual usage, providing maximum flexibility and scalability. This approach is most relevant for use cases characterized by **variable costs** and **spiky demand profiles**, such as GenAI tools involving per-token charges (L.E.K. Consulting, 2025; Zuora, n.d.).

2.5.2.1 Token-Based Pricing

For fundamental AI providers (the underlying platforms hosting LLMs), **tokens** have emerged as the standard pricing metric (Metronome, 2025).

- **Definition and Function:** Tokens serve as an **abstraction layer** that links the granular consumption of computational resources (like GPU hours or memory) to the customer's user engagement, making the complex underlying technology easier for the customer to understand.
- **Structure:** Pricing is typically differentiated based on the type of token consumed. OpenAI's GPT-4o, for instance, charges separate rates for input tokens ($2.50 per 1 million) and output tokens ($10.00 per 1 million).

- **Strategy:** Token pricing allows the provider to scale easily from individual developers to large enterprises. It also allows AI providers to potentially retain more value even if their underlying systems become more cost-effective over time.

2.5.2.2 Output Units and Activity Metrics

For application-layer SaaS vendors embedding AI features, pricing metrics often shift away from tokens (which are opaque to the end user) towards units that are **tangible and correlated with measurable customer value**:

- **Output Units:** Pricing tied to the creation or volume of an output, such as per-minute video generation (Synthesia) or per-minute of transcribed meeting content (Fireflies.ai) (L.E.K. Consulting, 2025; Metronome, 2025).
- **Activity-Based Metrics:** These relate pricing to the completion of specific, valued workflows. Examples include charging per sales development workflow completed (e.g., leads identified) rather than simply charging for every thousand emails sent (Khanna et al., 2025). Decagon, for instance, prices AI agents per conversation or per resolution (Metronome, 2025).

Choosing the right meter is crucial; it must **correlate with perceived value**, be easy for the customer to understand, and must not create disincentives that discourage feature adoption. Activity-based metrics, while easier to implement and forecast than pure outcome models, are a practical starting point toward capturing business value (Bain & Company, 2025; Khanna et al., 2025).

2.5.2.3 Prepaid Credits and Committed Spend

To mitigate the revenue volatility inherent in pure consumption (Pay-As-You-Go) models while addressing customer predictability concerns, most UBP is enacted through **prepaid credits** or **committed spend contracts**.

- **Prepaid Credits (PayGo):** Customers purchase credits upfront, often with a low minimum commitment (e.g., OpenAI's PayGo model). This lowers the adoption barrier but protects the vendor by instantly cutting off service when credits run out, preventing revenue loss(Metronome, 2025).
- **Committed Spend:** This enterprise-focused structure requires customers to pre-commit to a minimum usage level (e.g., capacity, compute) in exchange for better pricing discounts. This provides the vendor with predictable recurring revenue and gives customers the necessary budget predictability (Zuora, n.d.).

2.5.3 Outcome-Based Pricing (OBP): The Strategic Ideal

Outcome-Based Pricing (OBP) represents the strategic long-term goal of AI monetization, as it achieves the maximal alignment between the price paid and the tangible, verifiable business results delivered by the AI feature (BCG, 2025; McKinsey & Company, 2025; Zuora, 2025).

2.5.3.1 Structure and Examples

OBP links charges directly to success metrics, ensuring customers pay only for realized benefits (L.E.K. Consulting, 2025).

- **Service Automation Outcomes:** Charging based on measurable service results. Examples include Zendesk and Intercom's Fin charging per successfully **resolved ticket or customer interaction**. Intercom's Fin charges $0.99 per resolution, which aligns costs directly with the value derived (a resolved query) (L.E.K. Consulting, 2025).
- **Business Impact Outcomes:** Linking charges to measurable business improvements, such as per qualified lead identified or based on increased crop yields (Coherent Solutions, 2025; Khanna et al., 2025).

2.5.3.2 Execution Challenges and Customer Resistance

Despite its conceptual appeal and potential for maximizing value capture, OBP remains **rarely adopted** in practice. The successful execution of OBP at scale is subject to several difficult constraints:

- **Attribution Complexity:** It is difficult to prove that the product alone caused the measurable outcome, especially in complex B2B environments where multiple systems are involved. For example, Zendesk had to develop a complex, seven-step flowchart in monthslong trials to verify successful resolution for billing purposes (Coherent Solutions, 2025; Khanna et al., 2025).
- **Customer Preference:** Enterprise buyers are often **uncomfortable** with OBP and prefer predictable consumption models. One AI SDR company found that 90% of customers opted for activity-based pricing over outcome-based pricing (per qualified lead) because activity-based pricing offered simplicity and better spend forecasting (Khanna et al., 2025; Metronome, 2025).
- **Time Lag:** Success confirmation must happen quickly; if the lag time between product execution and measurable outcome is too long, the model is unworkable (Khanna et al., 2025).

Because of these difficulties, companies often resort to offering flexible contracts (e.g., Intercom offers annual credit buckets alongside OBP options) or hybrid models that provide free trials to prove value *before* committing to variable charges (McKinsey & Company, 2025; Metronome, 2025).

2.5.4 Operationalizing Monetization: Billing, Agility, and GTM

The adoption of dynamic AI monetization strategies necessitates a company-wide transformation, treating pricing not merely as a downstream function but as a central, cross-functional business system.

2.5.4.1 Re-architecting Billing and Infrastructure

The shift to usage-based and hybrid models inevitably **breaks existing billing systems** that were built for fixed-fee, seat-based subscriptions (Metronome, 2025).

- **The Fragmentation Problem:** Most incumbent SaaS companies run separate GTM stacks (Sales-Led Growth using CPQ/ERP; Product-Led Growth using Stripe), neither of which cleanly supports real-time usage pricing. Pricing flexibility is fundamentally **limited by billing capability**.
- **Telemetry Mandate:** Successful deployment requires investing in **robust product telemetry** (real-time usage tracking) to support billing accuracy and transparency. Companies often lack this capability when starting the transition (Bain & Company, 2025).
- **Enabling Tools:** This need has spurred the growth of specialized monetization platforms (like **Metronome**, Maxio, Stripe, and Orb) that help companies ingest real-time usage events, automate complex billing logic, and unify contract management across different systems. OpenAI, for instance, adopted Metronome specifically to manage its self-serve PayGo and complex enterprise contract requirements effectively and scalably (L.E.K. Consulting, 2025).

2.5.4.2 Pricing Agility and Market Responsiveness

In the competitive and rapidly evolving AI market, speed and adaptability in pricing are critical. The cost of LLM inference is declining quickly (reportedly by over 80% per year over the last two years) (Khanna et al., 2025). Pricing strategies must be agile to remain competitive and ensure that models stay accessible to users (Metronome, 2025).

- **Frequent Updates:** Successful players like OpenAI regularly adjust their pricing to align with new model launches and fluctuating usage costs.
- **Execution Precision:** Rolling out pricing updates across a diverse customer base (from solo developers to bespoke enterprise contracts) requires **precise control and orchestration** across backend systems to avoid errors and preserve customer confidence. Billing systems must be capable of propagating changes (to APIs, dashboards, and invoices) instantly.

2.5.4.3 Go-to-Market (GTM) Transformation

Sales and customer success teams must undergo significant transformation to sell variable AI features effectively.

- **The Sales Enablement Gap:** Sales representatives are typically trained and compensated to sell stable, predictable contracts (seats) and struggle to sell variable usage or outcome metrics (Bain & Company, 2025). This leads them to default to selling flat rates, undermining the hybrid strategy.
- **New Capabilities:** Sales teams require **new tools**, such as ROI calculators and granular usage dashboards, to forecast variable consumption for customers and justify the premium AI price by linking usage directly to measurable outcomes (Bain & Company, 2025).
- **Shifting Mindset:** Compensation and incentive structures must shift to reward consumption and long-term expansion rather than just upfront deal size. Successful companies reframe the value proposition to focus on productivity gains and cost savings compared to human labor (e.g., "Here's what a resolution would cost if a human did it—we're cheaper") (Bain & Company, 2025; Metronome, 2025).

The overall trend shows that companies that provide buyers with **clear expectations** (via caps, committed spend, or flat rates) are the ones that successfully unlock usage and expansion, reinforcing the idea that **predictability, not just the lowest price point, drives enterprise AI adoption**.

2.6 Identified Deficiencies in Existing Literature, Study Contribution, and Conclusion

The preceding review demonstrates that the infusion of Artificial Intelligence (AI) and Generative AI (GenAI) capabilities into Software as a Service (SaaS) products has created a **structural misalignment** that renders traditional pricing and business models inadequate. Despite the urgency of AI monetization, universally regarded as a **"board-level imperative"** (Metronome, 2025), systematic academic research documenting and analyzing the practical strategies adopted by firms remains nascent and fragmented. This section consolidates the critical deficiencies in the existing body of work, clearly articulates the precise contribution of the proposed study, and concludes the literature review.

2.6.1 Identified Deficiencies in Existing Literature

The shortcomings in current academic and industry literature fall into distinct thematic categories, justifying the need for a rigorous empirical investigation:

1. Insufficiency of Traditional Pricing Models and Cost Estimation Research: The foundational economic logic underpinning decades of SaaS pricing assumed **high fixed development costs and near-zero marginal distribution expenses**. This assumption is invalidated by embedded AI, which

introduces **significant and variable marginal costs** that scale linearly with usage, primarily through per-token or per-inference charges.

Prior cost estimation literature, such as empirical validations of models like SLIM, COCOMO, and Function Points, focused entirely on measuring **fixed, upfront costs** like man-months of effort and delivered source instructions (DSI/KSLOC). These models, while useful for managing traditional software development, were designed to predict costs for components characterized by predictable investment curves. Such research is fundamentally **insufficient** for the AI era because it fails to address the financial risk posed by variable, transactional compute costs (e.g., per-token API calls). The discontinuity between the established low-marginal-cost paradigm and the new high-variable-cost reality of AI leaves a severe theoretical gap in pricing strategy guidance.

2. Lack of Empirical Classification and Strategic Rationales: While industry reports descriptively identify emerging monetization avenues—such as usage-based metering, token-based pricing, and complex hybrid structures—these insights are largely anecdotal, fragmented, and **lacking empirical grounding**. There is a pronounced deficiency in systematic documentation of the **prevalence** and **classification** of these models across different market segments.

Crucially, the literature seldom delves into the **strategic and behavioral rationales**—the "why"—behind executive pricing decisions. Executives must navigate profound trade-offs: balancing the necessity of consumption-based pricing to cover variable costs against **customer anxiety over unpredictable bills**. Current research does not adequately document how management negotiates these tensions, such as choosing between maximizing long-term revenue via predictable contracts (to reduce churn risk) and maintaining margins via usage metering.

3. Absence of Segment Specificity and Outcome Linkage: Most analyses of AI monetization strategies focus broadly on large, established technology firms (like OpenAI or Salesforce) or on general market trends. There is a critical research vacuum concerning the strategies adopted by the **mid-sized B2B SaaS segment (US$1M–$100M ARR)**. This segment is vital as it is sufficiently mature to have established pricing models but agile enough to be actively experimenting with AI monetization, making it an ideal environment for observing emerging best practices.

Furthermore, empirical studies **linking pricing model choice to key performance metrics** (such as customer retention, revenue expansion, gross margin, or Average Revenue Per User [ARPU]) are scarce. Without this linkage, practitioners lack objective evidence to determine which monetization strategy successfully supports long-term growth and profitability in the AI context.

4. Neglect of Operational, Ethical, and Regional Dynamics: The shift to consumption-based models reveals internal organizational weaknesses, requiring new **telemetry, billing, and financial infrastructure** that many traditional firms lack. The literature needs empirical exploration of how companies are addressing this **billing fragmentation** and the associated **sales enablement gap**.

Finally, there is limited research exploring **cross-regional variation** (e.g., comparing US strategies with India's booming SaaS ecosystem), nor is there sufficient empirical engagement with the inherent **ethical complexities** of algorithmic pricing, specifically concerning fairness, transparency (the "black box" problem), and accountability. These systemic issues compound the monetization challenge by introducing potential discriminatory outcomes and creating information asymmetries between sellers and buyers.

2.6.2 Positioning the Proposed Study and Contribution

The proposed **mixed-methods empirical investigation** is designed specifically to address these gaps, transforming descriptive observations into actionable, empirically validated knowledge.

1. Holistic Methodological Rigor: The study utilizes a mixed-methods design, combining **quantitative analysis of over 100 pricing pages** with **qualitative interviews with 3 to 5 senior executives**. This approach is crucial because quantitative data captures the market prevalence ("the what"), while qualitative interviews uncover the strategic and behavioral constraints and motivations ("the why"). This depth allows the research to move beyond merely listing models to explaining the complex trade-offs executives face between driving adoption, maximizing revenue, and managing unpredictable costs.

2. Targeted and Actionable Findings: By concentrating exclusively on the **mid-sized B2B SaaS segment ($1M–$100M ARR)**, the research delivers findings that are highly relevant to this underserved yet high-impact segment, providing urgently needed benchmarks. The classification of prevalent monetization models (e.g., hybrid, usage-based, outcome-based, subscription add-ons) will provide a clear map of industry practices.

3. Theoretical and Practical Advancement: Academically, this study bridges the gap between technology adoption and business model innovation, establishing new insights into **value creation and capture** in the context of intelligent systems that violate traditional cost assumptions. Practically, the research addresses a **critical operational need** by systematically revealing which pricing models are most common and providing executives with the strategic rationales that drive success. The research will culminate in a **practical, evidence-based decision-making framework**. This outcome directly

supports the broader success and financial sustainability of the SaaS industry by helping firms avoid pitfalls like margin traps and customer churn when commercializing expensive AI features.

2.6.3 Conclusion

The contemporary SaaS industry stands at an inflection point, necessitated by the profound **structural misalignment** between its legacy fixed-cost economics and the variable-cost dynamics of AI-embedded features. The market has universally recognized that **per-seat models are, in some cases, obsolete**, driving a rapid and necessary convergence toward hybrid and usage-based models. Hybrid consumption models, combining the predictability of recurring subscriptions with metered usage for cost-intensive AI features, are emerging as the dominant strategy to balance vendor viability and customer demand for price predictability.

The challenge of monetizing AI features is thus fundamentally a problem of strategic governance and execution, requiring a continuous balancing act among cost recovery, adoption incentives, and demonstrable customer value. This research contributes by moving the conversation beyond descriptive anecdote to provide the **empirical evidence and strategic rationales** required for sustainable growth in the AI+SaaS era. The framework resulting from this study will serve as a vital guide for leaders seeking to navigate this complex commercial frontier.

Chapter 3: Research Methodology

3.1 Introduction

This chapter details the research methodology employed to conduct a mixed-methods empirical investigation into the pricing and packaging strategies for Artificial Intelligence (AI) features adopted by Software as a Service (SaaS) companies. The primary objective of this study is to systematically map the prevalent monetization models within a defined market segment and to uncover the strategic rationales that inform executive decision-making in this domain.

The current discourse on AI monetization is dominated by anecdotal evidence, practitioner commentary, and high-level industry reports. While valuable, this body of knowledge lacks the systematic, empirical grounding necessary to provide reliable benchmarks or validated frameworks for industry leaders. The strategic importance of a robust methodology, therefore, cannot be overstated. It provides the structured pathway required to move beyond conjecture and build an evidence-based understanding of the complex interplay between the new cost structures of AI, evolving customer value perceptions, and competitive market dynamics. A systematic and transparent research process is essential to ensure that the findings are credible, verifiable, and capable of making a substantive contribution to both academic theory and management practice.

This chapter lays out a roadmap of the research process. It begins by articulating the research philosophy that underpins the entire study, establishing the paradigmatic foundation for the subsequent methodological choices. It then details the mixed-methods sequential explanatory research design, justifying its selection as the most appropriate approach to address the study's objectives. Following this, the chapter outlines the sampling strategies for both the quantitative and qualitative phases, the specific data collection procedures, and the corresponding data analysis techniques. The chapter concludes by addressing the measures for ensuring validity and trustworthiness, discussing the ethical considerations that governed the study, and acknowledging the inherent limitations of the chosen methodology. The chapter therefore begins with an exposition of the research philosophy, as this foundational paradigm is the intellectual cornerstone that dictates the choice of research design and all subsequent methodological procedures.

3.2 Research Philosophy and Approach

This study adopts a pragmatist research philosophy, appropriate for research that seeks to address complex, real-world managerial problems rather than test a single theoretical paradigm (Saunders et al.,

2019). The objective of the research is not theory verification in isolation, but the development of an empirically grounded and practically useful framework for AI monetization decisions in SaaS firms.

The study follows an abductive research approach(Dubois & Gadde, 2002). Existing literature on SaaS pricing, platform monetization, and usage-based pricing provides an initial theoretical sensitization that informs data collection and coding. Empirical observations from the quantitative analysis of SaaS pricing structures and qualitative executive interviews are then used to iteratively refine, extend, and contextualize these theoretical constructs. This abductive logic enables the integration of established pricing theory with emerging industry practices related to AI-driven cost and value dynamics, culminating in a decision-oriented framework grounded in both theory and observed practice.

3.3 Research Design

The research design provides the logical blueprint that connects the research questions to the empirical evidence and, ultimately, to the study's conclusions. It specifies the sequencing and integration of data collection and analysis procedures to ensure that the evidence gathered is appropriate for addressing the research problem in a coherent and systematic manner.

To achieve the stated research objectives, this study adopts a mixed-methods sequential explanatory design. This design consists of two distinct and consecutive phases, beginning with quantitative data collection and analysis, followed by qualitative data collection and analysis. The findings from the quantitative phase inform the focus and structure of the subsequent qualitative phase (Tashakkori & Teddlie, 2010).

The selection of a sequential explanatory design is deliberately aligned with the study's primary objective of not only documenting AI monetization strategies in SaaS firms, but explaining the strategic rationales underlying them. The initial quantitative phase provides a structured overview of current practices by systematically analysing pricing and packaging information from over 100 SaaS company websites. This phase establishes the empirical landscape by identifying dominant monetization models, measuring their prevalence across firm segments, and revealing observable patterns and variations. In doing so, it addresses the "what" and "how many" dimensions of the research problem.

These quantitative findings then inform the second, qualitative phase. Insights from the initial mapping guide the formulation of targeted interview questions and sampling focus, enabling deeper exploration of the strategic reasoning, trade-offs, and contextual constraints that shape observed pricing decisions. The qualitative phase therefore addresses the "why" and "how" questions, providing explanatory depth that cannot be derived from quantitative analysis alone. This sequential structure supports an abductive

research logic, whereby empirical patterns observed in the data are iteratively interpreted and refined in light of existing theory and executive sense-making.

The sequence and purpose of each phase are as follows:

• **Phase 1 (Quantitative):** Systematically map the prevalence and characteristics of AI monetization models across the target SaaS segment, addressing the "what" and "how many" questions.
• **Phase 2 (Qualitative):** Explore the strategic reasoning, challenges, and contextual factors underlying the observed pricing models through in-depth executive interviews, addressing the "why" and "how" questions.

In summary, the mixed-methods sequential explanatory design enables the integration of quantitative breadth with qualitative depth. By combining large-sample descriptive analysis with focused executive insight, the design provides a robust foundation for developing a practical, evidence-based framework for AI monetization decision-making in SaaS firms.

3.4 Sampling Strategy

A sampling strategy is the process of select ing a representative subset of individuals or cases from a larger population to draw conclusions about that population. A well-defined sampling strategy is critical for ensuring the relevance and credibility of the research findings. Given the mixed-methods design of this study, distinct sampling strategies are employed for the quantitative and qualitative phases to align with their respective objectives.

The **target population** for this research consists of global Software as a Service (SaaS) companies. To ensure the findings are targeted and actionable, the **sampling frame** is further narrowed to companies with an estimated annual recurring revenue (ARR) between $1 million and $100 million USD. This specific revenue range was deliberately chosen to create a focused and highly relevant sample. It excludes early-stage startups that may not yet have stable or publicly articulated pricing models, as well as large enterprise-level corporations whose pricing structures are often opaque, custom-negotiated, and not publicly available. Where direct and recent revenue numbers are not available, the revenue will be estimated based on the number of employees listed on LinkedIn.

3.4.1 Quantitative Sampling Strategy

The sampling method for the quantitative phase is a **non-probability, purposive sampling technique**. This approach is appropriate as the objective is not to achieve statistical generalization to the entire global population of SaaS companies, but rather to construct a rich, descriptive dataset of companies that meet a specific and relevant set of criteria.

The process of identifying candidate firms will involve a systematic review of established industry databases and reports, such as **Crunchbase, Latka, SaaS 1000 listings, and venture funding reports**. Each potential company will be vetted against a strict set of inclusion and exclusion criteria to ensure the final sample is highly relevant and consistent. The criteria is detailed in the below.

Table 1: Inclusion and Exclusion Criteria

Inclusion Criteria	**Exclusion Criteria**
Is a Software as a Service company.	Companies outside the specified ARR range (1M-100M).
Has an estimated ARR between $1M and $100M USD.	Companies with "Contact Us for Pricing" as their only option.
Provides publicly accessible pricing and packaging information on their website.	Agencies or consulting firms.
Offers distinct, identifiable AI-powered features.	

The target sample size for this quantitative phase is **100+** qualifying SaaS companies. This size is considered sufficient to provide a robust sample for descriptive statistical analysis, allowing for the identification of significant patterns, the calculation of model prevalence, and a comparative analysis across the defined revenue sub-segments ($1-10M; $10-50M; $50-100M).

3.4.2 Qualitative Sampling Strategy

For the qualitative phase, a **purposive sampling** method will also be employed. In this context, the purpose is to deliberately select participants who possess deep, relevant, and direct knowledge of the phenomenon under investigation. This method ensures that the interview data is rich with the specific insights required to explain the strategic rationale behind AI monetization decisions. Participants will be recruited through professional networks, targeted outreach on platforms such as LinkedIn, and referrals.

The inclusion and exclusion criteria for participants are designed to identify individuals with direct strategic authority and experience.

- **Inclusion:** Must be a senior-level executive (e.g., CEO, CPO, CRO, VP of Product/Marketing) at a SaaS company meeting the quantitative phase criteria, with direct involvement in their company's AI monetization strategy.
- **Exclusion:** Individuals in roles without strategic pricing authority or from companies outside the study's scope.

The target sample size for the qualitative phase is **3 to 5** executives. This smaller sample size is appropriate for qualitative inquiry, where the primary goal is not statistical generalizability but rather depth and richness of information. The aim is to achieve **thematic saturation**, the point at which conducting further interviews is unlikely to generate new or significant themes related to the research questions.

3.5 Data Collection Methods

The data collection methods are the specific procedures and tools used to gather the information required to answer the research questions. The choice of methods is directly linked to the research design and the nature of the data needed for each phase of the study. This research employs two distinct methods: a systematic analysis of public documents for the quantitative phase and semi-structured interviews for the qualitative phase.

3.5.1 Quantitative Data Collection: Pricing Page Analysis

The primary method for quantitative data collection is a systematic analysis of publicly available information from the websites of the 100+ companies in the sample. This process involves a methodical review and data extraction from company pricing pages such as pricing, product feature lists, frequently asked questions (FAQs), and related information. This direct approach ensures that the data accurately reflects the company's publicly stated pricing and packaging strategy.

To ensure consistency, reliability, and objectivity in the data collection process, a predefined codebook will be developed and systematically applied. The codebook will function as a structured coding instrument, explicitly defining each variable, its operational meaning, and the allowable categories for classification (Krippendorff, 2018). This approach ensures that all company pricing pages are evaluated against identical criteria, thereby reducing researcher bias and improving the reliability and replicability of the dataset.

The codebook will specify both **observational variables** and **classified variables**, with clear decision rules for each. The key variables recorded for each company will include:

- **Pricing structure** (observed and classified), such as subscription, usage-based, tiered, or hybrid models
- **Primary value metric** used for pricing (observed and classified), such as per user, per account, per token, per credit, or per outcome
- **Number of publicly visible pricing tiers or plans**
- **Presence, absence, and explicit labeling of AI-related features**, based solely on stated terminology
- **List prices** for any publicly disclosed plans
- **Disclosure of usage caps, limits, or overage rules**, if explicitly stated

All variables will be coded conservatively. Where information is not explicitly disclosed or cannot be clearly classified using predefined rules, it will be recorded as "Not disclosed" rather than inferred.

3.5.2 Qualitative Data Collection: Semi-Structured Interviews

The primary method for qualitative data collection is **semi-structured interviews** with senior executives from a subset of the sampled companies. This method was chosen as the most effective means of gathering rich, in-depth data on the strategic thinking and contextual factors that influence AI monetization decisions.

The semi-structured format provides an optimal balance between consistency and flexibility. It utilizes a qualitative questionnaire, which ensures that a core set of key topics is covered with each participant. At the same time, the format provides the flexibility for the researcher to ask follow-up questions, probe deeper into emergent themes, and allow participants to elaborate on insights they deem important. This adaptability is crucial for uncovering the nuanced rationales that a purely structured interview would miss (Kvale & Brinkmann, 2015).

The interviews will be conducted by letting the participants fill in the questionnaire as per their convenience. The qualitative questionnaire provides the thematic structure for the interview guide, covering three main areas: Company and Product Context, AI Feature Development and Monetization, and the Competitive and Market Landscape. Procedurally, each interview will be recorded, contingent upon receiving explicit participant consent. Following the interview, the recordings will be transcribed verbatim to create a rich, accurate textual dataset that serves as the foundation for the subsequent thematic analysis.

3.6 Data Analysis Techniques

Data analysis is the process of inspecting, cleansing, transforming, and modeling data with the objective of discovering useful information, informing conclusions, and supporting decision-making. The data analysis plan for this study is directly aligned with its mixed-methods sequential explanatory design, employing distinct techniques for the quantitative and qualitative datasets before integrating the findings to produce a holistic understanding.

3.6.1 Quantitative Data Analysis

The quantitative data collected from the analysis of pricing pages will be coded and organized within a structured database, with subsequent analysis conducted using appropriate statistical software (e.g., Microsoft Excel for descriptive summaries and SPSS for cross-tabulations). The primary analytical technique employed will be **descriptive statistics**. This approach is ideal for summarizing and organizing the coded data to reveal patterns, frequencies, and central tendencies within the sample.

Specific outputs generated from this analysis will include:

- **Frequency distributions:** These will be used to determine the prevalence of different monetization models and value metrics (e.g., the percentage of firms using subscription add-ons versus credit-based systems).
- **Cross-tabulations:** This technique will be used to explore relationships between variables, specifically to analyze differences in pricing approaches across the defined company sub-segments based on annual revenue ($1-10M; $10-50M; $50-100M).
- **Measures of central tendency:** Metrics such as the mean and median will be calculated for numerical variables like the number of pricing tiers offered, providing further descriptive insight into industry norms.

3.6.2 Qualitative Data Analysis

The qualitative data, derived from the verbatim transcripts of the executive interviews, will be analyzed using **thematic analysis**. This well-established method is a process for identifying, analyzing, and reporting patterns (themes) within data. This study follows the systematic, 6 phase framework developed by Braun & Clarke in 2006, widely recognized as an important guide for conducting good thematic analysis.

The six phases of the analytical process are as follows:

- **Familiarization with the data:** Reading and re-reading the interview transcripts to gain a deep and holistic understanding of the content.

- **Generating initial codes:** Systematically identifying interesting features of the data and assigning descriptive codes to relevant text segments.
- **Searching for themes:** Collating the initial codes into potential overarching themes and identifying broader patterns of meaning.
- **Reviewing themes:** Checking if the themes are coherent and work in relation to both the coded extracts and the entire dataset.
- **Defining and naming themes:** Ongoing analysis to refine the specifics of each theme, articulating its scope and essence.
- **Producing the report:** The final analysis, involving the selection of compelling extract examples and relating the analysis back to the research question.

3.6.3 Integration of Quantitative and Qualitative Data

The critical final step in the analysis process is the integration of the findings from both the quantitative and qualitative phases. In keeping with the sequential explanatory design, this integration will occur primarily at the interpretation and reporting stage. The quantitative findings will provide the broad structural overview, while the qualitative findings will provide the explanatory depth and context. For example, a quantitative finding that "X% of companies use a hybrid model" will be explained and enriched by qualitative themes detailing *why* executives chose that model, the trade-offs they considered (e.g., balancing predictability and cost recovery), and the challenges they faced in implementation. This synthesis of "what" and "why" is essential to fulfilling the primary research objective of developing a practical, evidence-based framework for executive decision-making.

3.7 Validity and Reliability / Trustworthiness

Establishing the quality and rigor of the research is paramount in any academic inquiry. In mixed-methods research, different criteria are used to assess the quality of the quantitative and qualitative components. For quantitative research, the key criteria are **validity** (accuracy) and **reliability** (consistency). For qualitative research, the equivalent concept is **trustworthiness**, which encompasses a set of criteria to establish the believability of the findings.

3.7.1 Quantitative Phase: Validity and Reliability

- **Reliability:** Reliability will be ensured through the development and consistent application of a detailed **codebook** for the pricing page analysis. This instrument provides explicit definitions for all variables and clear instructions for coding, ensuring that the data collection process is consistent across all 100+ cases and would yield similar results if repeated.

- **Validity: Face validity** and **content validity** are addressed by collecting data directly from the primary source: the official pricing pages of the sampled companies. This ensures that the data accurately reflects the publicly stated monetization strategies of these firms, providing a direct measure of the phenomenon under investigation.

3.7.2 Qualitative Phase: Trustworthiness

The trustworthiness of the qualitative findings will be established by addressing its four core criteria:

- **Credibility (Confidence in the truth of the findings):** Credibility will be enhanced through **triangulation**, by comparing findings from the executive interviews with data from the quantitative analysis and existing industry reports. For instance, credibility is substantially enhanced if a quantitative finding—such as a high prevalence of hybrid credit-and-subscription models—is directly explained by a qualitative theme from executive interviews detailing the strategic trade-off between cost recovery and customer budget predictability. This convergence of evidence from distinct data sources validates the findings more powerfully than either source could alone.
- **Transferability (Applicability in other contexts):** Transferability will be supported by providing a **thick description** of the participants' organizational contexts and the research setting, allowing readers to make their own informed judgments about the applicability of the findings to their situations.
- **Dependability (Showing that the findings are consistent and could be repeated):** Dependability will be addressed by maintaining a clear and detailed **audit trail**. This involves meticulously documenting the research process, including data collection procedures, analysis decisions, and thematic development.
- **Confirmability (A degree of neutrality or the extent to which the findings are shaped by the respondents and not researcher bias):** Confirmability will be achieved through the audit trail and the practice of **reflexivity**, where the researcher acknowledges and reflects upon their own potential biases and assumptions throughout the research process.

3.8 Ethical Considerations

This research will be conducted in strict accordance with the ethical guidelines of the IIBM Institute of Business Management. The integrity of the research process and the welfare of the participants are of utmost importance. The primary ethical considerations relate to the qualitative phase involving human participants (Bell et al., 2019).

The following specific ethical procedures will be implemented:

- **Informed Consent:** All participants in the qualitative phase will be provided with a clear explanation of the study's purpose, procedures, and their role. Written consent will be obtained before conducting any interviews, confirming their voluntary participation.
- **Confidentiality and Anonymity:** The identity of participants and their companies will be kept strictly confidential in the final thesis. All data will be anonymized, and responses will be used solely for academic purposes.
- **Data Storage and Security:** All collected data, including interview recordings and transcripts, will be stored securely on password-protected devices to prevent unauthorized access.
- **Voluntary Participation:** Participants will be informed that their involvement is entirely voluntary and that they have the right to withdraw from the study at any time without penalty.

3.9 Limitations of the Methodology

A hallmark of rigorous academic research is a transparent acknowledgment of its inherent limitations. While the chosen mixed-methods design is robust and well-suited to the research objectives, it is important to recognize the boundaries that shape the scope and interpretation of the findings. The primary limitations are as follows.

- **Sampling Bias:** The use of non-probability sampling means the findings may not be statistically generalizable to the entire SaaS population. The exclusion of companies with opaque "Contact Us for Pricing" models may also skew the sample towards certain types of companies.
- **Geographic Scope:** While the sample is global, specific regional analyses (e.g., US vs. India) will be exploratory rather than conclusive due to sample size constraints.
- **Data Scope:** The study focuses on the commercial strategies for pricing AI and does not evaluate the technical performance or efficacy of the AI features themselves.
- **Qualitative Sample Size:** The small sample size for interviews, while appropriate for achieving thematic saturation, limits the generalizability of the qualitative findings.

3.10 Chapter Summary

This chapter has provided a detailed and comprehensive account of the research methodology designed to investigate the monetization of AI features in the SaaS industry. The methodological framework is anchored in a pragmatic research philosophy, which justifies the use of an integrated, problem-focused approach. The study employs a mixed-methods sequential explanatory design, beginning with a broad

quantitative analysis of over 100 company pricing pages, followed by in-depth qualitative interviews with senior executives to explain and enrich the quantitative findings.

The chapter has detailed the purposive sampling strategies for each phase, the specific procedures for data collection through pricing page analysis and semi-structured interviews, and the corresponding techniques for data analysis, namely descriptive statistics and thematic analysis. Furthermore, it has outlined the rigorous measures that will be taken to ensure the validity, reliability, and trustworthiness of the findings, and has affirmed the study's commitment to the highest ethical standards. Finally, the inherent limitations of the methodology have been transparently acknowledged. This robust and well-justified methodology provides a systematic and rigorous plan to address the research objectives, ensuring that the conclusions drawn are grounded in credible, empirical evidence. The application of this methodology will yield the findings and results that are presented and discussed in the subsequent chapters.

Chapter 4. Research Results

4.1 Introduction

This chapter presents findings from the quantitative and qualitative phases of the study in sequence, consistent with a sequential mixed-methods design (Creswell & Plano Clark, 2018). The primary objective of this chapter is to report the empirical results regarding how mid-sized B2B Software as a Service (SaaS) firms monetise embedded Artificial Intelligence (AI) features, strictly presenting observed data without theoretical interpretation or causal explanation.

To achieve a comprehensive understanding of current AI monetisation strategies, the chapter combines data from two distinct methodological phases. The quantitative phase consists of a systematic pricing-page analysis of 509 global SaaS companies, within which 293 firms were identified as actively offering AI features. Consistent with the scope of this study, the primary analytic sample focuses on firms generating between $1 million and $100 million in Annual Recurring Revenue (ARR). Following the establishment of these baseline metrics, the qualitative phase draws upon semi-structured interviews with senior SaaS executives to capture the strategic considerations, operational trade-offs, and packaging decisions shaping these observed pricing structures.

The chapter is structured to logically transition from broad market realities to specific executive experiences. Section 4.2 presents the quantitative results, detailing the prevalence of various AI monetisation models such as pure subscription, usage-based, credit-based, and hybrid structures alongside cross-segment analyses based on revenue tiers and geographic regions. Section 4.3 then details the qualitative findings, outlining the emergent themes derived from executive interviews regarding the operational and market tensions of pricing AI, such as managing cost uncertainty and navigating competitive pressures. Section 4.4 integrates these two datasets, demonstrating the empirical alignment between the quantitative pricing patterns and the qualitative themes without introducing theoretical explanation. Finally, Section 4.5 provides a summary of the established patterns.

In accordance with the study's analytical framework, this chapter focuses exclusively on reporting the integrated findings. Ultimately, this chapter answers *what* pricing and packaging strategies are currently observed in the market and *what* executives report doing in practice, rather than interpreting *why* these strategies are optimal. The theoretical interpretation and strategic implications of these findings are reserved for discussion in Chapter 5.

4.2 Quantitative Results

This section presents the findings from the quantitative pricing-page analysis, establishing the empirical market reality of how B2B Software as a Service (SaaS) companies are currently monetising and packaging embedded Artificial Intelligence (AI) features. The data strictly reflects observable, public-facing commercial strategies, detailing the structural prevalence of specific models without attempting to deduce the internal decision-making processes that led to these configurations.

4.2.1 Sample Characteristics

The quantitative phase of this study involved a systematic review of the public-facing pricing and packaging pages of 371 global SaaS companies. The primary objective of this data collection was to identify the baseline prevalence of AI integration and the corresponding commercial models utilized by these firms. Across the total dataset of 371 companies, 215 firms (representing exactly 58.0% of the overall sample) were identified as actively offering AI-powered features or enhancements within their product suites.

To ensure the analysis precisely aligned with the study's predefined scope, the companies were categorised by Annual Recurring Revenue (ARR) tiers. The entire analytic sample of 371 firms was systematically filtered to fall within the $1 million to $100 million ARR spectrum, explicitly excluding early-stage startups engaged in initial price discovery and large enterprise incumbents whose pricing heavily relies on custom, opaque contracts. The final revenue tier distribution within this refined sample is as follows: 87 companies fall within the $1 million to $10 million tier; 207 companies constitute the $10 million to $50 million tier; and 77 companies belong to the $50 million to $100 million tier.

Furthermore, instances where pricing variables or structural details were obscured, gated behind sales contact forms, or unlisted on public pages were systematically recorded as "UNKNOWN" in the dataset. These omissions are noted throughout the following sections as data limitations inherent to the empirical reality of public SaaS pricing pages.

4.2.2 Prevalence of AI Monetisation Models

An analysis of the commercial structures applied to AI features reveals a heterogeneous landscape, characterised by a mix of traditional SaaS billing logic and emergent consumption-based frameworks. Within the observed dataset, pure subscription models remain the overwhelmingly dominant monetisation structure. Specifically, 191 occurrences of subscription-only pricing were observed. This

indicates that a significant majority of the market continues to monetise AI by embedding it within recurring, fixed-fee access plans rather than metering granular AI consumption.

Despite the dominance of pure subscriptions, a measurable structural transition toward hybrid monetisation models is empirically evident. The data indicates that 67 companies have adopted a "Subscription + Usage-based" model, where a baseline recurring fee is coupled with variable charges tied to specific consumption limits. Similarly, 20 companies utilize a "Subscription + Credit-based" structure, allocating a defined number of AI credits within the base subscription and establishing a framework for additional credit purchases. Purely variable models are notably less common: 11 companies rely entirely on pure usage-based pricing, and only one company utilises a standalone credit-based model. Additionally, the public-facing monetisation structure for 81 companies could not be definitively identified and was recorded as UNKNOWN.

The specific metrics used to meter AI value further illustrate the fragmentation and lack of standardisation in the market. Among the firms that explicitly define an AI usage metric, "Credits" emerged as the most frequent unit of measurement (23 instances), followed by traditional user "Seats" (16 instances). Beyond these, companies deploy a wide array of highly granular, action-oriented consumption metrics. These include "Requests" (10 instances), "Generations" (6 instances), "Minutes" (5 instances), and "Tokens" (3 instances). A long tail of highly specific, product-centric metrics was also recorded, with single instances of monetisation tied to "Projects," "Occurrences," "Words," "Queries," "Tasks," "Headshots," "Resolutions," "Cases," "Monthly Active Users (MAU)," and "Tests". It is critical to note that for the vast majority of the sample—298 entries—the primary AI metric was not explicitly identifiable on the public pricing page.

4.2.3 AI Packaging Strategies

Packaging strategies—defined as how AI features are bundled, tiered, or separated from core SaaS offerings—exhibit distinct structural patterns across the dataset. The quantitative analysis reveals that tiered packaging is the most widely adopted approach, observed in 157 companies. In this model, AI capabilities are mapped across progressive plan levels (e.g., Basic, Pro, Enterprise), acting as a feature differentiation lever across the pricing spectrum.

Hybrid packaging, which combines elements of different strategies (such as offering basic AI features in a standard tier while isolating advanced AI capabilities as an independent add-on), is the second most prevalent approach, utilized by 76 companies. This indicates a bifurcated deployment strategy

where foundational AI value is treated as a core product expectation, while specialized AI tools are separated to capture additional revenue.

Additionally, 52 companies rely on Custom or Enterprise packaging, requiring prospective buyers to engage directly with sales teams for AI access and pricing. Freemium models, where basic AI functionality is provided at no cost prior to hitting a paywall, were observed in 23 companies. Conversely, only 3 companies in the sample utilized a single-plan packaging structure, highlighting that “one-size-fits-all” packaging is highly uncommon in AI-enhanced SaaS. As with the monetisation metrics, packaging strategies for 60 companies could not be categorized based on publicly available data.

4.2.4 Revenue-Based Comparisons

A cross-segment analysis across the three target revenue bands reveals observable structural shifts in AI monetisation models as firm maturity increases.

Table 2: Revenue-Based Monetization Models

Revenue Range	Subscription	UNKNOWN	Subscription + Usage-based	Usage-based	Subscription + Credit-based	Credit-based
$10M–$50M	115	42	38	6	6	0
$1M–$10M	38	26	11	2	9	1
$50M–$100M	38	13	18	3	5	0

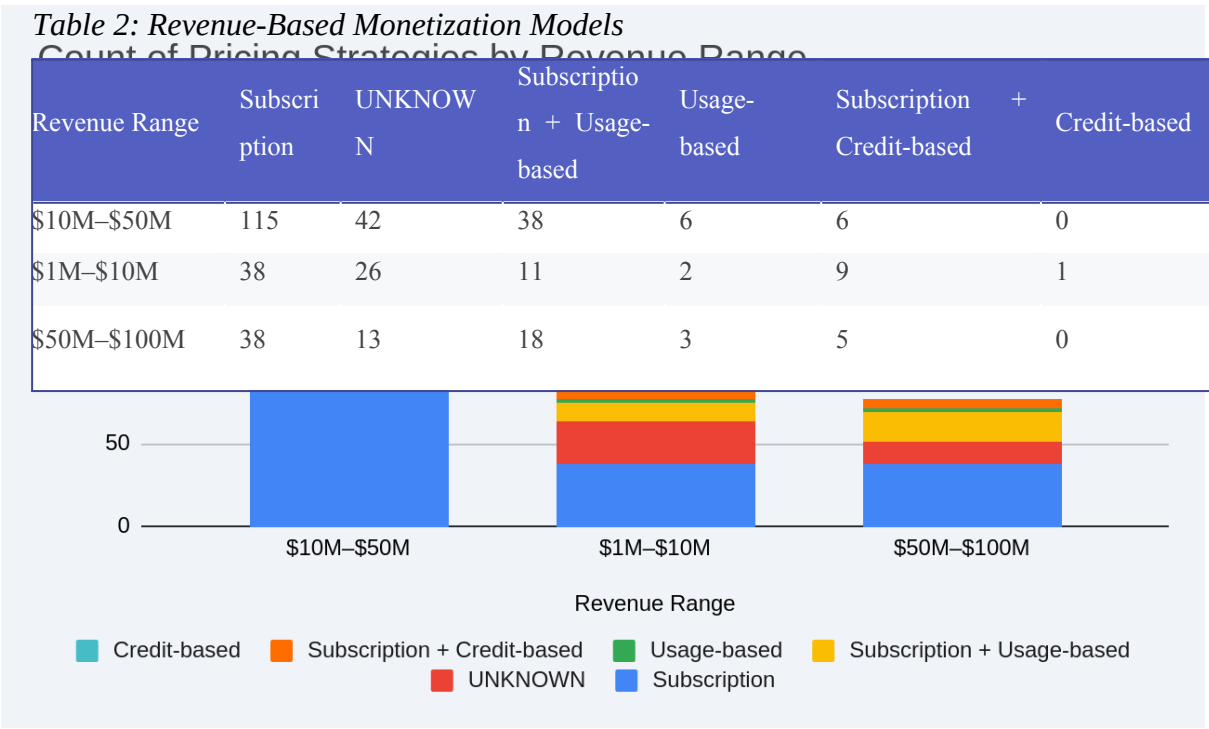


Figure 2: Pricing Strategy by Revenue

In the **$1 million to $10 million ARR tier** (87 companies), pure subscription models are the most frequent structure, accounting for 38 firms. Early-stage hybridisation is present: 11 companies use a Subscription + Usage-based model, and 9 use a Subscription + Credit-based model. Pure usage-based (2 companies) and pure credit-based (1 company) models are marginally represented. Notably, this segment has a relatively high proportion of "UNKNOWN" structures (26 companies), reflecting potential ongoing price discovery.

In the **$10 million to $50 million ARR tier** (207 companies), the largest segment of the analytic sample, the pure subscription model scales up significantly, observed in 115 companies. Concurrently, the deployment of the Subscription + Usage-based approach is utilised by 38 firms. Subscription +

Credit-based models are used by 6 firms, while pure usage-based pricing maintains a minor presence at 6 firms. The monetisation structure for 42 companies in this tier remains UNKNOWN.

In the **$50 million to $100 million ARR tier** (77 companies), the pure subscription model remains the leading choice (38 companies). However, the Subscription + Usage-based model exhibits a proportionally strong presence, utilized by 18 companies. Subscription + Credit-based approaches (5 companies) and pure usage-based models (3 companies) are also represented, alongside 13 companies with UNKNOWN structures.

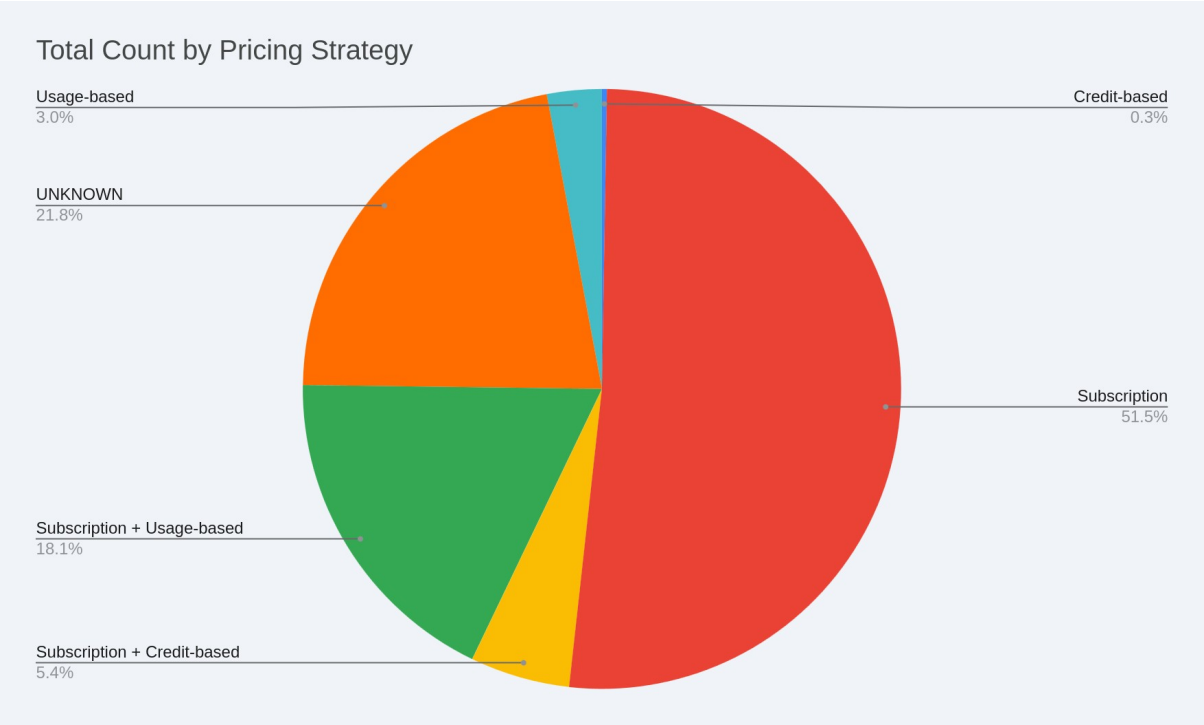


Figure 3: Pricing Strategies

Table 3: Pricing Strategies

Pricing Strategy	Total_Count
Credit-based	1
Subscription	191
Subscription + Credit-based	20
Subscription + Usage-based	67
UNKNOWN	81
Usage-based	11

Comparing across the bands, the data demonstrates that while pure subscriptions remain the dominant baseline across all segments, reliance on "Subscription + Usage-based" structures is consistently observable and scales proportionally as companies mature into the higher revenue tiers. Conversely, pure usage-based and credit-based models without a subscription baseline remain consistently rare regardless of company size.

4.2.5 Regional Comparisons

The final dimension of the quantitative analysis sought to identify observable differences in monetisation models across geographical boundaries, specifically comparing established markets like the US with emerging SaaS ecosystems such as India.

An empirical assessment of the pricing pages reveals that public-facing commercial structures are heavily standardised for a global audience. The quantitative dataset did not exhibit distinct, structurally segregated pricing models based on the regional origin of the software provider. The application of tiered packaging, the statistical dominance of pure subscriptions, and the emergent use of Subscription + Usage-based models remained structurally uniform across the sampled websites. Furthermore, pricing metrics were uniformly presented without regional differentiation, and the vast majority of firms listed standard public pricing in US Dollars regardless of their headquarters' location.

Consequently, the quantitative data establishes that, from an external market positioning standpoint, mid-sized B2B SaaS firms present a homogenized global approach to AI monetisation. Measurable variances in regional adoption, localized discounting strategies, or market readiness are not transparently codified on standard public pricing pages. The nuanced regional differences that influence these homogenous public structures must therefore be explored through the qualitative experiences of the executives managing these models.

4.3 Qualitative Results

This section presents the findings from the qualitative phase of the study, detailing the strategic considerations, operational trade-offs, and market realities that shape how mid-sized B2B SaaS firms monetise embedded AI capabilities. Interview transcripts were analysed using thematic analysis following the procedures outlined by Braun and Clarke (2006). To ensure analytical rigor, iterative coding cycles and systematic pattern identification were employed (Miles, Huberman & Saldaña, 2014). Furthermore, steps were taken to enhance the credibility and consistency of the coding process throughout the analysis (Lincoln & Guba, 1985). Data collection continued until a sufficient depth of

empirical observation was achieved, aligning with established principles of data saturation (Guest et al., 2006).

4.3.1 Overview of Interview Sample

The qualitative dataset comprises responses from senior executives and product leaders representing mid-sized B2B SaaS organisations. The sampled firms operate across a diverse array of industry verticals, including enterprise resource planning (ERP), real estate and property management, customer data analytics, digital retail receipts, financial operations, portfolio management, and user experience (UX) validation.

To protect the commercial sensitivities of the pricing strategies discussed, the participating firms and executives have been anonymised in this reporting. Respondents are instead identified by their primary industry vertical or functional focus (e.g., "an executive at a financial operations platform"). The participating firms primarily target mid-market to enterprise-level customers, aligning precisely with the $1 million to $100 million ARR parameter established in the quantitative phase. The qualitative data bridges directly to the quantitative archetypes by capturing the internal decision-making processes such as cost management, customer feedback, and competitive benchmarking that ultimately manifest as the public-facing subscription, usage-based, and hybrid models identified in Section 4.2.

4.3.2 Emergent Qualitative Themes

The thematic analysis of the executive responses revealed three distinct, recurring patterns regarding AI monetisation. These themes highlight the practical tensions SaaS leaders navigate when packaging and pricing AI features within their existing product ecosystems.

Theme 1: Managing AI Cost Uncertainty

The most prominent theme across the qualitative data is the persistent internal tension between driving customer adoption of new AI features and managing the unpredictable infrastructure costs associated with large language models (LLMs). Executives frequently cited the direct computational expense of AI infrastructure as a primary, restrictive factor shaping their packaging and monetisation decisions.

For example, an executive from a digital receipts and customer engagement platform noted that their central pricing considerations revolved heavily around "managing AI infrastructure and development costs". This firm explicitly embedded AI as a core capability rather than an optional add-on to accelerate adoption, but they acknowledged the ongoing challenge of continuously improving

capabilities without compromising product stability or gross margins. Similarly, a leader from a portfolio management solution highlighted that "AI infra consumption" has become a critical decision parameter, noting a "fundamental change in consumption patterns" regarding product value.

To mitigate margin compression and control financial risk, several firms reported implementing strict operational guardrails. An executive from a financial operations platform serving enterprise and bank customers detailed a bifurcated approach to cost management. They reported that standard "deterministic models" are explicitly chargeable, whereas "probabilistic models" are currently provided for free but are strictly "guardrailed" to specific point use cases. By restricting LLM usage to customised, trained models operating only on customer-specific data, the firm intentionally limits its compute exposure.

In contrast, other executives reported temporarily absorbing AI costs to achieve market penetration. A founder of a vertical SaaS platform providing retail analytics observed that "initially our AI costs will be higher as customers won't pay more than a certain amount". Their strategy relies on the assumption that the firm must simply "bear the costs of AI" in the short term to solve the customer's problem at scale, with the expectation of normalizing data ingestion costs into user pricing at a later stage. Across these varied approaches, the qualitative data consistently demonstrates that the fear of unpredictable variable costs fundamentally limits the widespread deployment of unmetered AI features.

Theme 2: Customer Resistance to Pure Usage-Based Pricing

A second major theme emergent from the data is widespread customer resistance to complex, purely usage-based pricing models. Despite the variable cost structure of AI compute, executives reported that B2B buyers strongly prefer predictable billing and often struggle to comprehend highly granular consumption metrics (such as tokens or compute minutes).

An executive from a real estate and property management platform encapsulated this sentiment, emphasizing that they "focused on controlling costs, keeping prices easy to understand, [and] matching what customers are willing to pay". This respondent advised peers to "avoid over-complicating usage models early on," noting that keeping AI pricing straightforward directly enabled healthier adoption and clearer value-based sales conversations.

Furthermore, executives observed that customers are unwilling to pay additional usage fees for AI if the technology merely enhances basic workflows. A leader in the customer data analytics sector explained their firm's clear delineation: "If AI enhances basic workflows, we don’t charge. As this improves the ux of the product... If it brings more value and features, we charge as an add-on". This

executive explicitly noted that "usage and credit charges are difficult," leading the firm to utilise transactional credits exclusively for highly distinct, tangible outputs, such as AI-enhanced e-receipts.

This preference for simplicity aligns with a broader industry push toward pricing based on business outcomes or the specific "job to be done," rather than raw AI usage. The executive from the retail analytics platform noted that because they are a vertical SaaS provider, they "package it as a customer jobs to do and we price it for the job". Because the AI agent solves a highly specific planning problem, customers "don't mind paying for the same" as long as the pricing maps to the user or agent, rather than abstract compute usage. Similarly, a respondent from a managed operations service noted that "AI features monetize best when embedded directly into daily workflows... not as standalone tools users must 'figure out'". The overarching executive consensus is accurately captured by the customer analytics leader's advice: "Price for adoption now. Price for outcomes later".

Theme 3: Competitive Pressure and Feature Parity

The third empirical theme highlights the intense external market pressure compelling SaaS firms to include AI capabilities simply to maintain competitive parity, often severely restricting their ability to monetise these features as premium add-ons.

Multiple executives observed that both direct competitors and the broader consumer AI market have fundamentally altered B2B customer expectations. An executive from a qualitative UX validation platform explicitly cited this anchoring effect, stating: "In today's world, AI features are available very commonly. Especially with Gen AI modules able to function all aspects of work to a certain extent... both competitors and the general gen AI modules that are freely available have had a big influence". As a direct result of this influence, the firm currently does not charge separately for basic AI assistance, though they anticipate introducing credit-based models for "more unique features" in the future.

This sentiment was mirrored by the executive from the digital receipts platform, who stated that "broader market expectations have pushed us to treat AI as a baseline capability rather than a differentiator". This competitive reality reinforced their decision to embed AI features directly into the core product to avoid "adoption friction". Even when companies opt for add-on models, competitive benchmarking remains a primary driver. The real estate platform executive noted that "seeing how competitors price AI and what customers expect made us keep AI pricing straightforward and affordable".

Additionally, executives noted that while the pressure to include AI is universal, actual market readiness to adopt fully autonomous AI varies significantly by domain, further complicating

monetisation. The executive from the financial operations platform observed that while basic LLMs are used externally, enterprise adoption of AI in "deep banking" or treasury remains hesitant due to financial risk. Consequently, AI is primarily used as a supportive tool "to make things ready for their review," meaning monetisation remains anchored to traditional platform sturdiness rather than the AI features themselves. Ultimately, the data shows that competitive pressure forces the rapid deployment of AI, but customer hesitation and the ubiquity of free consumer AI tools frequently cap the premium that mid-sized B2B SaaS firms can command.

4.3.3 Summary of Qualitative Patterns

The qualitative interviews present a consistent empirical reality: mid-sized B2B SaaS firms are caught between the rising variable costs of AI infrastructure and a customer base that demands predictable, outcome-aligned pricing while simultaneously expecting AI as a standard, baseline feature.

The theme of managing cost uncertainty illustrates that internal infrastructure tracking is actively shaping external packaging, with firms adopting guardrails or eating early margins to prevent runaway compute expenses. Concurrently, the theme of customer resistance to usage-based pricing highlights that buyers reject the complexity of token or credit metering unless it is tied to highly specific, tangible tasks. Consequently, executives report a strong preference for embedding AI into daily workflows and pricing it based on user seats or specific "jobs to be done." Finally, the theme of competitive pressure reveals that the ubiquity of consumer GenAI tools and competitor parity heavily anchors customer willingness to pay, forcing many firms to offer foundational AI capabilities at no additional charge to prevent adoption friction. Together, these three themes comprehensively map the operational and market tensions executives report experiencing as they attempt to commercialise embedded AI.

4.4 Integration of Quantitative and Qualitative Results

This section integrates the findings from the quantitative pricing-page analysis (Section 4.2) with the emergent themes from the qualitative executive interviews (Section 4.3). By mapping the observable market structures against the reported internal decision-making processes, the data demonstrates clear empirical alignment between how AI is publicly monetised and the operational tensions experienced by SaaS leaders.

The quantitative dominance of pure subscription models and tiered packaging directly aligns with the qualitative theme regarding customer resistance to usage-based pricing. The quantitative phase revealed that 191 companies utilise a subscription-only model for AI features, making it the most prevalent

structure across the sample. This observable market reality is mirrored by executive reports that B2B buyers strongly prefer predictable billing and frequently reject complex, granular consumption metrics. The qualitative data illustrates that to drive adoption, firms are actively avoiding pure usage-based models, which were quantitatively observed in only 11 companies, in favor of embedding AI value into familiar recurring subscriptions. The integration of these datasets confirms that the structural preference for fixed-fee or broadly tiered packaging corresponds directly with customer demands for commercial simplicity and outcome-aligned pricing.

Similarly, the emergence of hybrid pricing models in the quantitative dataset maps clearly to the qualitative theme of managing AI cost uncertainty. The quantitative analysis identified 87 companies employing either "Subscription + Usage-based" or "Subscription + Credit-based" structures, a trend that becomes proportionally more prevalent as firms mature through the $10 million to $100 million ARR tiers. Qualitatively, executives articulated a persistent internal tension between driving feature adoption and mitigating the unpredictable variable costs of underlying AI infrastructure. The hybrid models observed in the public market represent the external manifestation of the operational guardrails executives described implementing internally. By anchoring AI access to a base subscription while applying variable usage caps or credit limits, companies simultaneously satisfy the customer desire for a predictable baseline while protecting their own gross margins from unchecked compute consumption.

Finally, the integration reveals an alignment between the widespread integration of AI and the qualitative theme of competitive pressure. The quantitative phase established that 58% of the sampled firms have deployed AI features, with monetisation structures uniformly presented on a global scale and showing no distinct regional variations in public-facing pricing pages. This structural homogenisation aligns tightly with executive reports that intense competitive pressure and the anchoring effect of widely available consumer AI tools force companies to offer foundational AI as a baseline expectation. Just as executives noted the necessity of including basic AI to avoid adoption friction and maintain feature parity, the quantitative data reflects a market where AI is heavily bundled and standardized globally. Together, the datasets illustrate a landscape where external market expectations accelerate AI deployment and standardise public pricing, even as firms continue to navigate complex internal cost structures.

4.5 Summary

This chapter presented the integrated empirical findings from the quantitative and qualitative phases of the study, mapping how mid-sized B2B SaaS firms monetise embedded AI features. The quantitative pricing-page analysis of 371 companies established that while pure subscription models and tiered

packaging remain the dominant market structures, there is a clear, scale-dependent shift toward hybrid monetisation. As firms mature through the $10 million to $100 million ARR tiers, "Subscription + Usage-based" and "Subscription + Credit-based" models become increasingly prevalent, indicating a structural transition away from purely unmetered access to AI capabilities.

The qualitative phase, drawing on semi-structured interviews with SaaS executives, surfaced three emergent themes that shape these public-facing models. First, executives face significant internal tension regarding the management of unpredictable AI infrastructure costs, often implementing operational guardrails to protect gross margins. Second, despite the variable cost of AI compute, there is widespread customer resistance to pure usage-based pricing; buyers strongly prefer the commercial simplicity and predictability of subscriptions or outcome-aligned billing. Third, intense competitive pressure and the anchoring effect of consumer AI tools frequently compel firms to offer foundational AI capabilities as baseline expectations rather than premium differentiators.

The integration of these datasets empirically established that current AI monetisation strategies are defined by a delicate operational balancing act. The quantitative prevalence of hybrid pricing and core subscriptions directly mirrors the qualitative reality of executives attempting to shield both their internal margins from runaway compute costs and their customers from complex, unpredictable billing.

Chapter 5 Discussion

5.1 Introduction

The global Software as a Service (SaaS) industry is experiencing a profound structural shift driven by the rapid integration of Artificial Intelligence (AI). Historically, SaaS economic models were optimized around high fixed development costs and near-zero marginal costs of distribution, an architecture that supported predictable recurring revenue through traditional per-seat subscriptions. However, the embedding of active AI capabilities fundamentally disrupts this established equilibrium by introducing significant and variable marginal costs, primarily due to the complex infrastructure and computational expense of large language model (LLM) inference. This dynamic presents a central problem: SaaS companies are struggling to design pricing and packaging strategies that can effectively cover variable compute costs while simultaneously aligning with enterprise customers' demands for budget predictability. Consequently, navigating this fundamental cost-value misalignment has driven a major pricing transformation trend across the software ecosystem.

The primary objective of this study was to address this structural challenge through a rigorous empirical investigation into the AI pricing and packaging models adopted by B2B SaaS companies operating within the $1 million to $100 million Annual Recurring Revenue (ARR) segment. To capture both the prevalence and the rationale behind emerging strategies, the research employed a sequential mixed-methods design. The study first conducted a quantitative pricing-page classification of mid-sized SaaS firms to map the objective market distribution of pure subscription, usage-based, credit-based, and hybrid monetization models. Building upon these structural baselines, qualitative interviews with senior SaaS executives were utilized to uncover the underlying strategic constraints, competitive pressures, and operational trade-offs that drive executive decision-making in practice.

The purpose of this chapter is to reconnect these empirical findings to the study's initial problem statement and broader pricing theory. It deliberately transitions the analysis from documenting **"what exists"** in the current market landscape to interpreting **"what it means"** for the future of SaaS economics in the AI era.

To build this understanding, the chapter is organized sequentially:

- **Section 5.2** provides a comprehensive summary of the key findings, strictly aligned with the core research objectives.

- **Section 5.3** serves as the intellectual core of the discussion, interpreting the empirical data through theoretical lenses such as platform economics, behavioral pricing theory, and risk-sharing dynamics to explain phenomena like the dominance of hybrid models and psychological price anchoring.
- **Section 5.4** translates these interpretations into actionable theoretical and managerial implications, presenting a structured framework for AI monetization decisions.
- Finally, **Sections 5.5 and 5.6** transparently outline the limitations of the study and propose high-value avenues for future research, before **Section 5.7** synthesizes the ultimate conclusion: that AI monetization is not merely a tactical pricing adjustment, but a structural business model evolution.

5.2 Summary of Key Findings

This section consolidates the empirical results derived from the mixed-methods investigation of 371 mid-sized B2B SaaS firms, anchoring these observed market realities against prior literature. The findings explicitly answer how companies are currently structuring their AI commercial models and the operational drivers behind these configurations.

5.2.1 Classification of AI Monetization Models

The quantitative analysis of SaaS pricing pages revealed a highly heterogeneous monetization landscape characterized by a mixture of traditional and emergent billing structures, reflecting a broader trend of pricing complexity growth in the industry. Out of the 371 companies analyzed, **215 firms (58.0%) actively offered AI features**.

Despite industry discourse heavily favoring usage-based and hybrid pricing, **pure subscription models remain overwhelmingly dominant, observed in 191 companies**. In these models, AI capabilities are typically embedded within recurring, fixed-fee plans rather than metered. However, a measurable structural transition toward hybrid models is evident: **67 companies utilize a "Subscription + Usage-based" framework, and 20 employ a "Subscription + Credit-based" model**. Pure consumption models are exceedingly rare, with only 11 companies relying entirely on pure usage-based pricing.

Regarding packaging, **tiered strategies are the most widely adopted (157 companies)**, where AI capabilities differentiate progressive plan levels. Hybrid packaging, which separates advanced AI tools as independent add-ons while bundling basic AI, is utilized by 76 firms. Metrics used to meter value

remain highly fragmented, with "Credits" (23 instances) and "Seats" (16 instances) being the most common identifiable units, followed by granular metrics like requests, generations, and tokens.

5.2.2 Prevalence by Revenue Segment

A cross-segment analysis of the $1 million to $100 million ARR spectrum demonstrates a scale-dependent evolution in pricing strategy.

- **$1M–$10M Tier (87 firms):** Pure subscriptions lead (38 firms), with early-stage hybridization present (11 Subscription + Usage-based, 9 Subscription + Credit-based) and a high proportion of "UNKNOWN" pricing structures (26 firms) indicating ongoing price discovery.
- **$10M–$50M Tier (207 firms):** Subscriptions scale significantly (115 firms), while the deployment of Subscription + Usage-based approaches expands to 38 firms.
- **$50M–$100M Tier (77 firms):** While subscriptions remain the baseline (38 firms), **the proportion of hybrid structures (18 Subscription + Usage-based) becomes structurally stronger**.

The data confirms that as companies mature into higher revenue segments, their reliance on complex hybrid monetization scales proportionally, mirroring broader theories of SaaS maturity and revenue scaling.

5.2.3 Geographic Differences

The quantitative dataset did not reveal structural segregation based on the regional origin of the software provider; instead, mid-sized SaaS firms present a highly homogenized, globalized approach to public-facing AI pricing. Metrics, tiered packaging, and the dominance of subscriptions remained uniform across regions, typically priced in US Dollars.

However, qualitative insights indicated nuances in cross-market SaaS evolution. Executives noted that **regional maturity impacts willingness to adopt and pay**. Customers in more developed or highly competitive regions are generally more open to trying and paying for AI features. Conversely, in emerging segments or highly regulated sectors (such as "deep banking" in India or other global markets), enterprise adoption remains hesitant due to financial risk and a preference for proven platform sturdiness.

5.2.4 Executive Decision Drivers

Thematic analysis of executive interviews highlighted three primary operational drivers shaping these commercial models.

- **Managing Cost Uncertainty:** Executives face immense pressure to manage the unpredictable infrastructure costs of AI. Firms frequently restrict unmetered AI features to protect gross margins, creating a tension between adoption growth and cost recovery.
- **Customer Predictability Demands:** There is widespread customer resistance to granular, pure usage-based pricing. B2B buyers fear unpredictable bills and struggle to map AI metrics to human workflows, forcing vendors to keep pricing straightforward to avoid adoption friction.
- **Competitive Pressure:** The ubiquity of free consumer AI modules and rapid competitor feature parity heavily anchor customer willingness to pay. Consequently, market expectations frequently compel SaaS firms to bundle foundational AI capabilities as baseline requirements rather than premium differentiators to prevent churn.

5.2.5 Cost–Value Alignment Strategies

To resolve the structural tension between variable AI infrastructure costs and customer demands for predictable billing, firms are deploying specific cost-value alignment strategies. Rather than charging strictly for compute, companies are attempting to price based on the "job to be done" or tangible business outcomes. For example, one executive noted successfully charging transactional credits for highly distinct outputs like AI-enhanced e-receipts, while embedding workflow enhancements for free. Another firm implemented strict operational guardrails to balance financial exposure: **deterministic AI models are explicitly chargeable, while probabilistic Large Language Models (LLMs) are provided for free but strictly guardrailed** to operate only on customer-specific data for point use cases. Ultimately, the overarching executive consensus is to "price for adoption now, price for outcomes later," absorbing initial variable costs to prove value before attempting complex, metered monetization.

5.3 Interpretation of Findings

This section serves as the intellectual core of the discussion, synthesizing the empirical market realities observed in the quantitative and qualitative data with the broader theoretical literature surrounding software economics. By intentionally transitioning the analysis from identifying *what* pricing strategies exist to theoretically interpreting *why* these specific configurations have emerged, this section explains how mid-sized B2B SaaS firms are actively resolving the fundamental economic, operational, and psychological tensions introduced by embedded Artificial Intelligence.

5.3.1 AI as a Structural Break in SaaS Economics

The empirical findings of this study provide robust confirmation that the integration of generative AI is not merely an incremental feature enhancement, but a profound **structural break** in the established economic architecture of Software as a Service. To fully understand the magnitude of this disruption, it must be contextualized against historical software economics. Traditionally, SaaS business models were predicated on an established economic logic characterized by massive, upfront fixed development costs and near-zero marginal costs of distribution. Classical software cost estimation literature historically relied on algorithmic models such as COCOMO (focused on Source Lines of Code) or Function Points, designed entirely to predict the fixed human effort (e.g., man-months) required for software development. Under this legacy paradigm, once the software was built, delivering the product to an additional user incurred negligible expense, allowing vendors to profitably scale via predictable, fixed-fee subscription tiers.

The quantitative discovery that exactly 58.0% of the sampled mid-sized firms (215 out of 371) now actively offer AI features illustrates the rapid obsolescence of this traditional low-marginal-cost paradigm. Embedded AI transforms software from a passive tool into an active, computational collaborator, introducing **significant and variable marginal costs** that scale linearly with user activity. The computational expense of running Large Language Models (LLMs) for inference is strictly transactional, frequently priced by foundational providers on a per-token or per-inference basis.

This creates a severe paradox for SaaS economics: whereas historically high product engagement signaled successful adoption and improved profitability, high AI inference volumes can now trigger immediate margin compression or even financial losses if the pricing strategy is misaligned with these variable costs. The qualitative data vividly confirmed this structural tension; executives across diverse sectors repeatedly cited "AI infra consumption" and "managing AI infrastructure and development costs" as the primary, restrictive factors shaping their product positioning. Because SaaS firms can no

longer rely on Kemerer's traditional algorithmic cost models to forecast their financial exposure, they are forced to completely restructure how they capture value, transitioning away from static value capture toward highly dynamic, consumption-aware business models.

5.3.2 Why Hybrid Models Dominate

Given this structural break, standard economic theory might suggest a total industry pivot toward pure usage-based pricing, as this would perfectly align the vendor's variable token costs with customer revenue. However, the empirical data reveals a remarkably different reality: pure usage-based models are exceedingly rare, observed in only 11 companies within the entire dataset. Instead, pure subscriptions remain the baseline (191 firms), while **hybrid pricing models** specifically "Subscription + Usage-based" (67 firms) and "Subscription + Credit-based" (20 firms), have emerged as the definitive, growing structural equilibrium, particularly as firms mature into the $50M–$100M ARR tier.

Through the theoretical lens of platform economics and behavioral pricing, this hybridization can be interpreted not as an aggressive monetization tactic, but as a necessary **risk-sharing mechanism** engineered to resolve competing market demands. On one side of the transaction, the SaaS vendor faces severe financial exposure; as noted in the literature, GenAI gross profit margins typically run 10 to 30 percentage points lower than traditional software, necessitating cost recovery. On the other side, the enterprise customer exhibits a profound psychological resistance to pure usage-based billing. Enterprise buyers fundamentally reject the financial risk of uncapped consumption, fearing "surprise bills" and lacking the confidence to forecast variable spend.

The hybrid archetype functions as the perfect theoretical compromise. By anchoring AI access to a familiar, predictable base subscription, the vendor satisfies the customer's requirement for standard procurement workflows and budget stability. Simultaneously, by introducing a capped "bucket" of credits or specific usage thresholds for advanced features, the vendor implements critical **revenue smoothing logic** while strictly capping their downside computational risk. As one executive explicitly noted, usage and credit charges are notoriously difficult to implement, so firms rely on standard subscriptions for basic workflow enhancements and deploy transactional credits exclusively to protect against high-cost power users. Consequently, the hybrid model dominates because it effectively bridges the gap between the vendor's variable cost reality and the enterprise buyer's demand for absolute predictability.

5.3.3 Psychological Anchoring and Customer Perception

The pervasive customer resistance to granular usage metrics such as tokens, generations, or compute minutes must be critically interpreted through behavioral pricing theory and customer value perception. The qualitative dataset underscores that B2B buyers struggle to mentally map abstract computational costs to tangible business value. As a real estate software executive advised, vendors must "avoid over-complicating usage models early on," prioritizing simplicity to maintain healthy adoption.

This resistance is heavily driven by **psychological anchoring**. Current literature highlights that B2B buyer expectations are profoundly influenced by ubiquitous, low-cost consumer AI applications like ChatGPT Plus, which anchors their willingness to pay at artificially low, flat rates (e.g., ~$20/month). When mid-sized SaaS vendors attempt to charge complex, variable fees that accurately reflect true enterprise-grade infrastructure and data processing costs, they encounter massive adoption friction due to this anchored disconnect. Customers do not inherently value "smarter" backend software; they value outcomes. This explains why the quantitative data shows firms eschewing abstract "tokens" (only 3 instances observed) in favor of more tangible metrics like "credits" (23 instances) or embedding the cost directly into the "seat" (16 instances).

Furthermore, the introduction of algorithmic and usage-based pricing inevitably introduces complex ethical and perception issues regarding market fairness. As established in the literature, the "black box" nature of AI systems creates a distinct information asymmetry between the vendor and the buyer. Customers who do not understand how underlying tokens are consumed inherently distrust metered models, viewing them as opaque and potentially exploitative. To restore perceived fairness and trust, executives actively package AI as a solution to specific "jobs to be done" rather than raw compute power. The consensus derived from the qualitative data summarized by the directive to "price for adoption now, price for outcomes later" demonstrates that SaaS firms are deliberately absorbing the information asymmetry risk in the short term to build trust through predictable, flat-rate pricing.

5.3.4 Strategic Trade-offs Identified

Executives designing these monetization structures do not operate in a vacuum; they exist at the nexus of several conflicting operational and market pressures. Synthesizing the qualitative thematic analysis reveals four primary strategic tensions that dictate the ultimate design of an AI pricing model.

1. Adoption Growth vs. Risk of Margin Erosion: This is the most critical tension observed in the dataset. Firms must decide whether to prioritize rapid market penetration or immediate financial sustainability. Some organizations choose to aggressively pursue adoption by intentionally absorbing the high computational costs. For example, the founder of a vertical retail analytics platform explicitly

stated that they "just have to bear the costs of AI" initially, accepting lower margins to solve the customer's problem at scale with the hope of normalizing costs later. Conversely, firms highly sensitive to margin erosion deploy strict operational guardrails. An executive at a financial operations platform resolving deep banking workflows noted that they provide "probabilistic models" for free only within strictly confined, custom-trained guardrails to prevent runaway usage.

2. Margin Protection vs. Risk of Churn (Competitive Parity): While vendors wish to charge premiums to protect their margins, intense external market pressure forces their hand. The ubiquity of free GenAI modules means that AI is rapidly transitioning from a premium differentiator to a baseline expectation. An executive from a digital receipts platform noted that "broader market expectations have pushed us to treat AI as a baseline capability". To avoid churn and prevent adoption friction, firms are forced to embed foundational AI into core subscriptions for free, absorbing the cost entirely to maintain competitive parity against rivals offering similar capabilities.

3. Commercial Simplicity vs. Cost Volatility: As established, customers demand simplicity, but simplicity exposes the vendor to cost volatility. If a firm offers unmetered AI access via a flat subscription (simplicity), they risk a small percentage of power users consuming massive amounts of compute, drastically undermining the product's unit economics. Executives resolve this tension through **bifurcated packaging**. As reported by a customer data analytics leader, if an AI feature merely "enhances basic workflows," the firm does not charge extra, viewing it as a UI improvement. However, if the feature generates a highly distinct, high-cost output (like an AI-enhanced e-receipt), it is strictly metered using transactional credits to manage the volatility.

4. Innovation Speed vs. Pricing Stability: The AI technology landscape is highly fluid, with the underlying costs of LLM inference declining rapidly. However, enterprise contracts require stability. Vendors struggle to lock in pricing for a technology whose capabilities and costs change monthly. The quantitative data reflects this uncertainty: 81 companies had an "UNKNOWN" overall pricing structure, and 298 companies obscured their specific AI metrics, indicating that a vast portion of the market remains in active price discovery. SaaS firms are resolving this by avoiding hard outcome-based guarantees early on, instead utilizing flexible "beta" periods or credit allocations that can be quietly adjusted on the backend without requiring a formal renegotiation of the master service agreement.

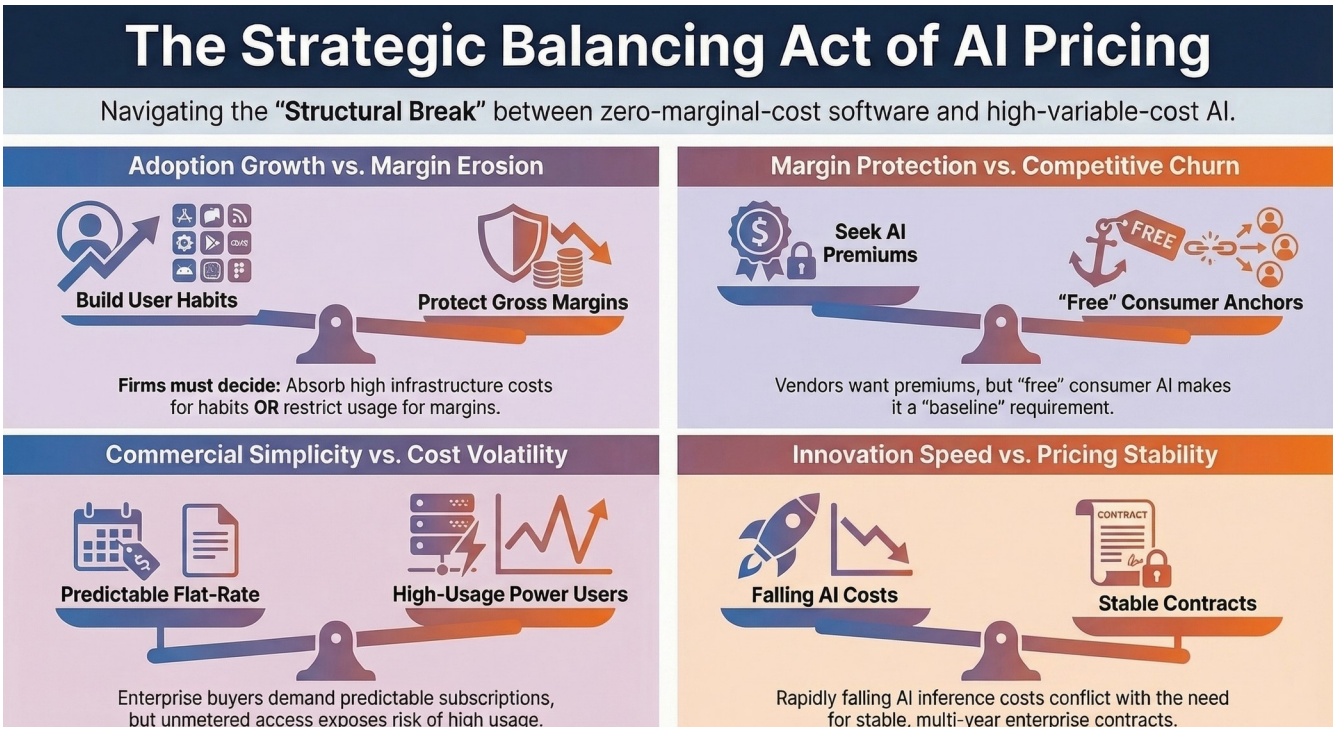


Figure 4: Strategic Trade-offs of AI Pricing

5.3.5 Revenue Segment Maturity Effects

The cross-segment analysis of the $1 million to $100 million ARR spectrum empirically confirms that AI monetization strategies are not uniform, but rather deeply dependent on an organization's structural capability and maturity. The quantitative dataset illustrates a clear, scale-dependent evolution in how firms package intelligence.

Smaller firms operating within the **$1 million to $10 million ARR tier** exhibit the highest rate of "UNKNOWN" pricing structures (26 out of 87 firms), indicating a market segment still heavily engaged in price discovery. These early-stage firms predominantly rely on pure, unmetered subscriptions (38 firms). Interpreted through the lens of organizational capability readiness, these smaller entities frequently lack the advanced product telemetry, robust billing infrastructure, and market leverage required to implement complex usage-based pricing. Furthermore, prioritizing market share over immediate margin, they often view AI as a loss-leader. As the founder of a vertical retail analytics platform in the dataset reported, they initially "just have to bear the costs of AI" to solve the customer's problem at scale, with the hope of normalizing data ingestion costs later.

Conversely, as firms mature into the **$50 million to $100 million ARR segment**, the monetization architecture undergoes a structural shift. In this tier, the proportion of hybrid structures specifically the

"Subscription + Usage-based" model (18 out of 77 firms) scales proportionally. This progression can be explained by organizational maturity theory: larger firms possess the financial discipline, specialized sales enablement, and technological infrastructure necessary to track granular consumption and enforce overage policies. By utilizing sophisticated monetization platforms (such as Metronome or Stripe) that can ingest real-time usage events, these mature firms successfully transition from "growth-first" bundling to the disciplined, risk-sharing hybrid models required to protect unit economics at scale.

5.3.6 Geographic Interpretation

A critical dimension of this research is the exploration of cross-regional variations in software pricing strategy, comparing established Western markets with emerging ecosystems like India's booming SaaS sector, which is projected to reach $50 billion ARR by 2030.

Quantitatively, the analysis of public pricing pages revealed a surprising **global homogenization**. Mid-sized B2B SaaS firms present structurally uniform public-facing pricing architectures regardless of their geographic origin, heavily relying on standardized tiered packaging and universally listing prices in US Dollars. This indicates that global SaaS markets have converged on a standardized external positioning language.

However, interpreting the qualitative data reveals profound underlying nuances regarding market readiness and regional ecosystem maturity. Executives explicitly noted that "geographies, industry sectors, [and] organisational maturity" drastically alter customer willingness to pay. Customers in highly developed, fiercely competitive markets (like the US) are increasingly open to trying and paying premiums for autonomous AI features to gain efficiency advantages. In contrast, in emerging domains or highly regulated regional sectors such as "deep banking" and enterprise operations in India enterprise buyers exhibit severe hesitation. In these markets, AI adoption carries higher perceived financial and operational risks, leading buyers to treat AI strictly as a supportive, probabilistic tool rather than a fully autonomous agent. Consequently, while Indian SaaS firms possess a distinct global "efficiency advantage" in software development, their ability to aggressively monetize AI locally remains anchored to traditional "platform sturdiness" and core business function reliability rather than the novelty of the AI features themselves.

5.3.7 Evolutionary Pathways

Synthesizing the empirical findings with the theoretical concepts of digital goods pricing allows for the proposition of a dynamic, **5-stage evolutionary pathway** for AI monetization. The data indicates that

SaaS firms do not select a static pricing model, but rather evolve sequentially as their internal technological capabilities and their customers' behavioral acceptance mature.

Stage 1: Bundled Free (Exploration): In an effort to overcome initial customer anxiety and respond to intense competitive pressure, foundational AI is embedded into core subscriptions at no extra cost. Firms explicitly "price for adoption now," absorbing the computational expense to train models, build user habits, and demonstrate clear value.

Stage 2: Premium Tier Gating (Differentiation): As capabilities improve, the firm shifts AI out of the foundational product and gates it behind higher-tier enterprise plans. This leverages AI to drive subscription upgrades while maintaining the predictable, flat-fee billing that enterprise procurement prefers.

Stage 3: Add-on Pricing (Value Isolation): Advanced, highly tangible AI workflows (such as "AI enhanced e-receipts" or discrete copilot tools) are isolated as distinct premium add-ons. This reflects a shift toward value-based pricing, capturing additional revenue for capabilities that clearly exceed standard product expectations.

Stage 4: Hybrid Usage-Credit Model (Risk Mitigation): To protect margins against power users, the firm introduces structural guardrails. Customers receive a capped "bucket" of pooled AI credits within their base subscription, with overages charged for excess consumption, perfectly balancing vendor risk with customer budget predictability.

Stage 5: Outcome-Based Experimentation (Value Alignment): The ultimate strategic ideal. The vendor moves entirely away from abstract metrics like tokens or seats, instead linking pricing strictly to verifiable, machine-executed business result such as per "resolved ticket" or "qualified lead".

While Outcome-Based Pricing (Stage 5) maximizes value capture and aligns perfectly with customer success, it remains exceedingly rare in the dataset due to complex attribution challenges and the persistent customer preference for simplified forecasting. Thus, the evolutionary pathway underscores that the transition toward sophisticated AI monetization is fundamentally constrained by the speed at which enterprise buyers develop trust in algorithmic outcomes.

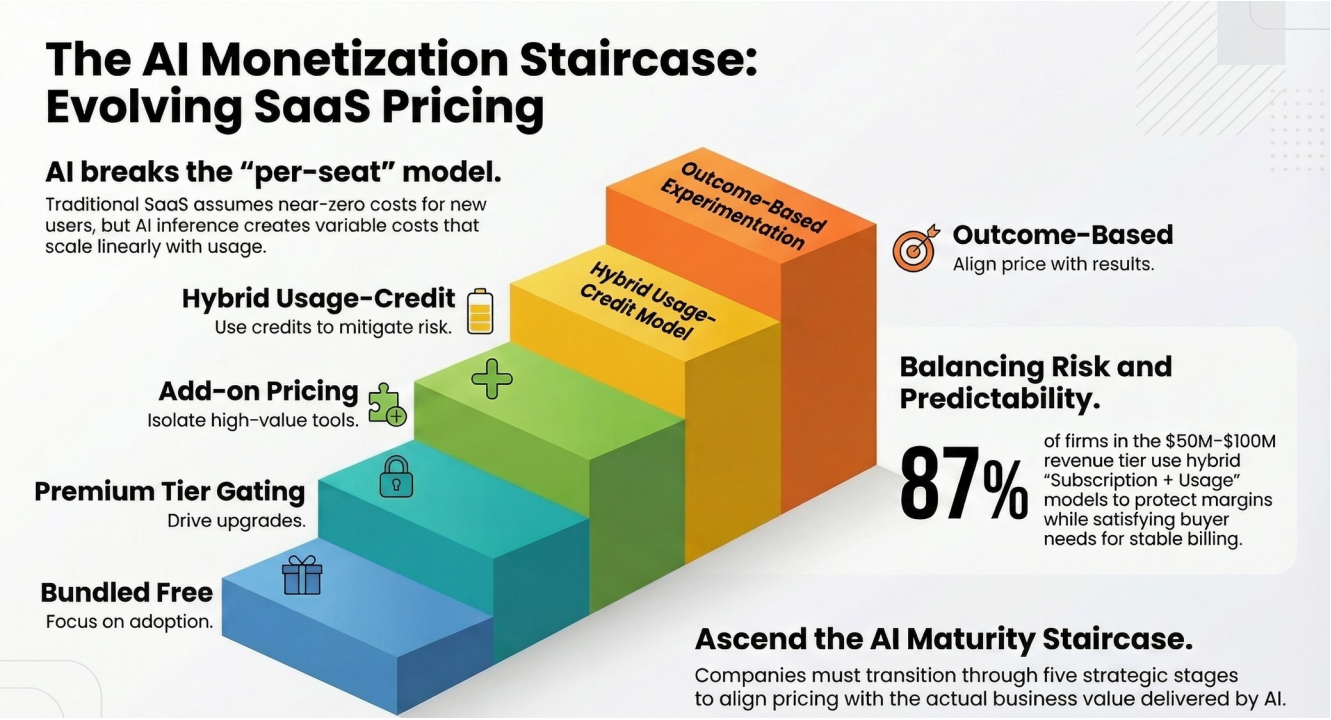


Figure 5: 5-stage evolutionary pathway for AI monetization

5.4 Implications

This section translates the empirical findings and theoretical interpretations into actionable insights. It details how the integration of embedded AI features structurally impacts academic frameworks of software economics, and provides a structured decision-making guide for SaaS practitioners navigating this commercial transition.

5.4.1 Theoretical Implications

This study formally advances the academic understanding of software economics by extending SaaS pricing theory into variable-cost environments. Historically, literature on software pricing and cost estimation has been rigidly anchored to a paradigm characterized by high upfront fixed development costs and near-zero marginal distribution expenses. However, the empirical results of this research confirm that the embedding of active AI features introduces significant, variable marginal costs that scale linearly with usage, primarily through LLM inference. By systematically documenting how mid-sized SaaS firms adapt their public-facing packaging to this structural break, this research successfully integrates dynamic AI cost structures into established pricing frameworks. It demonstrates that the variable nature of AI forces a fundamental departure from traditional, static value capture models, explicitly extending prior conceptualizations of dynamic software pricing (Esan, 2021).

Furthermore, the study provides robust empirical evidence supporting hybridization as the dominant emerging logic in SaaS monetization. While previous literature has explored the gradual growth of pricing complexity within digital markets, this research explicitly quantifies the structural shift toward risk-sharing mechanisms, noting that 23% of the sampled AI-enabled firms have already deployed "Subscription + Usage-based" or "Subscription + Credit-based" models. This finding directly addresses existing gaps in the literature on SaaS pricing heterogeneity by illustrating how firms actively resolve the tension between capturing variable value and satisfying enterprise demands for predictability (Saltan & Smolander, 2021; García-Fernández et al., 2025).

Finally, these findings enrich the broader digital goods pricing literature and business model innovation theory. Rather than viewing AI integration merely as an incremental feature enhancement, this study positions it as a profound technological shift that requires a complete restructuring of the SaaS business model. By demonstrating how companies transition from bundled loss-leaders to complex outcome-based metrics, the research grounds abstract theories of competence-enhancing innovation into observable, real-world commercial realities (Paschen et al., 2020).

5.4.2 Managerial Implications

The practical challenges of monetizing intelligent features represent a critical hurdle that has rapidly elevated from a routine product-level decision to a "board-level imperative" (Metronome, 2025). The empirical findings and qualitative executive insights translate into highly actionable guidance for product managers, pricing strategists, and executive boards navigating this transition.

A. Framework for AI Monetization Decisions To assist executives in designing sustainable, market-aligned pricing strategies, this study proposes a sequential 5-step decision-making framework:

- **Step 1: Assess AI Cost Volatility:** Organizations must strictly quantify their variable cost exposure, distinguishing between predictable "deterministic" operations and highly variable, compute-intensive "probabilistic" LLM queries.
- **Step 2: Define Strategic Objective (Adoption vs. Margin):** Leadership must align on a core objective whether to prioritize rapid market penetration by temporarily absorbing compute costs to establish user habits, or to prioritize immediate gross margin protection.
- **Step 3: Evaluate Customer Predictability Sensitivity:** Firms must measure their target buyer’s tolerance for variable billing. Because enterprise buyers typically reject the financial risk of uncapped consumption, vendors must determine the level of budget predictability required to prevent adoption friction.

- **Step 4: Select Pricing Archetype:** Based on the intersection of cost volatility and customer sensitivity, the firm must select the most appropriate structural packaging model.
- **Step 5: Build Transition Path:** Firms must recognize that AI pricing is dynamic. Leaders should actively map a transition path from early "bundled free" exploration to mature, outcome-based monetization as both product capabilities and customer trust evolve.



Figure 6: 5-Step Strategic Framework

B. Recommended Archetypes The empirical dataset validates five distinct architectural approaches to AI packaging that firms can deploy depending on their maturity:

- **Growth-first bundling:** Embedding AI features into core subscriptions at no extra charge to drive immediate adoption, maintain competitive parity, and train models.
- **Premium tier gating:** Restricting advanced AI capabilities exclusively to higher-tier enterprise plans, forcing subscription upgrades while maintaining predictable flat fees.
- **Credit-based hybrid:** Providing a fixed, pooled allowance of AI credits within a baseline subscription, generating overage revenue only when power users exceed their thresholds.
- **Metered usage with base subscription:** Applying strict consumption metrics (e.g., tokens, requests, generations) on top of a recurring platform fee for highly variable use cases.

- **Enterprise negotiated outcome model:** Linking pricing directly to verifiable business results (e.g., successful ticket resolutions), representing the ultimate strategic ideal for value alignment.

C. Risk Management Guidance SaaS leaders must proactively implement mechanisms to manage the intense financial risks of AI deployment. A primary directive is avoiding catastrophic "margin traps," where high product engagement paradoxically leads to unrecoverable financial losses due to underlying API costs (McKinsey, 2025; SaaStr / Unusual Ventures, 2025). Executives should deploy "soft guardrails," such as restricting unmetered probabilistic AI features to specific, low-risk use cases while explicitly metering heavy deterministic tasks. Simultaneously, to prevent adoption suppression caused by enterprise anxiety over "surprise bills," firms must abstract raw computational metrics behind tangible, value-aligned units that customers inherently understand, ensuring pricing remains comprehensible and predictable (Bain & Company, 2025).

D. Implications for Investors and Boards The shift to AI-integrated business models necessitates strong, top-down governance. Investors and boards must treat pricing agility as a central strategic lever, rather than a downstream consequence of software development (Bain & Company, 2022). This requires mandating structural investments in robust product telemetry and dynamic billing infrastructure capable of orchestrating complex usage events across fragmented enterprise systems. Furthermore, boards must establish strong pricing experimentation governance, ensuring that executive teams possess the technological and organizational flexibility to rapidly iterate their commercial models as underlying LLM inference costs decline and competitive market expectations mature.

5.5 Limitations of the Study

While this study provides robust empirical insights into how mid-sized B2B SaaS firms monetize embedded AI features, it is subject to several methodological and contextual constraints. These limitations are outlined transparently below; rather than invalidating the findings, they serve to define the boundaries of the study’s applicability and generalizability.

5.5.1 Sample Limitations

The quantitative phase of this research relied entirely on the systematic analysis of public-facing pricing and packaging pages. This approach inherently introduces a "public pricing page bias." While it successfully captures external market positioning, it obscures the reality of custom-negotiated enterprise contracts and off-book discounting. The empirical data itself highlights this limitation: the

pricing structures for 81 companies, packaging strategies for 60 companies, and the primary AI metrics for 298 companies could not be definitively identified and were recorded as "UNKNOWN". Furthermore, the qualitative phase drew upon a limited sample of senior executive interviews. While this depth of qualitative inquiry is excellent for uncovering strategic rationales and operational tensions, it restricts the broad generalizability of the specific internal practices discussed across the entire SaaS ecosystem (Braun & Clarke, 2006). Additionally, the deliberate scoping to firms with $1 million to $100 million ARR excludes early-stage startups engaged in pure price discovery and massive enterprise incumbents.

5.5.2 Geographic Scope Constraints

The quantitative dataset analyzed SaaS firms targeting a global audience, and the qualitative interviews focused heavily on perspectives from distinct ecosystems, primarily the United States and India. Although the quantitative findings indicated that public-facing pricing pages are heavily homogenized globally, the underlying strategic nuances and market readiness realities shared by executives may not perfectly map to other major economic regions, such as Europe, Southeast Asia, or Latin America. The profound heterogeneity of digital markets means that regional differences in regulatory environments, software purchasing power, and data privacy expectations naturally limit the universal generalization of these specific pricing archetypes (Saltan & Smolander, 2021).

5.5.3 Reliance on Self-Reported Performance Measures

A critical limitation of the qualitative findings is the reliance on self-reported executive perceptions regarding the success and impact of their AI monetization strategies. Statements concerning "healthy adoption," "margin control," or the prevention of customer churn were based entirely on the subjective assessments of the interviewed leaders. The study did not have access to audited financial data, internal product telemetry, or verified gross margin metrics to quantitatively validate these claims. Consequently, the reported outcomes represent perceived success rather than objective, financially audited performance.

5.5.4 Cross-Sectional Design

This study employed a cross-sectional design, capturing a specific "snapshot in time" of the AI monetization landscape. The artificial intelligence sector is characterized by unprecedented technological evolution and cost volatility. For instance, the cost of underlying LLM inference has been reported to decline by over 80% year-over-year. Because the cost dynamics driving these pricing

models are highly fluid, the specific pricing archetypes and thresholds observed in this dataset may undergo rapid alteration as foundational AI models become cheaper, more efficient, and more commoditized.

5.5.5 Classification Subjectivity

Finally, the quantitative classification of dynamic software pricing structures involves an inherent degree of subjectivity. The proliferation of complex, overlapping pricing categories such as distinguishing exactly where a "tiered subscription with usage caps" ends and a "subscription plus credit-based overage" begins, introduces coding interpretation risks. While predefined codebooks and systematic pattern identification were utilized to maintain rigor, the fragmented and unstandardized nature of AI pricing metrics across the industry leaves room for interpretative variance.

5.6 Recommendations for Future Research

The findings and limitations of this study expose critical gaps in the current understanding of AI monetization within digital markets, offering several high-value avenues for future academic inquiry. As the technological capabilities of AI and the corresponding cost structures continue to evolve rapidly, subsequent research must expand beyond cross-sectional observations to capture the long-term, dynamic nature of this transition.

5.6.1 Longitudinal AI Pricing Evolution Studies

Given the unprecedented technological evolution and cost volatility of underlying large language models, the current cross-sectional data represents a "snapshot in time". Future research should employ longitudinal methodologies to track specific SaaS cohorts over a 3-to-5-year period. Observing how firms dynamically adjust their pricing architectures as they move from early experimentation toward maturity will provide essential empirical data to map structural innovation pathways in software. Such studies could definitively confirm whether the proposed 5-stage transition model from "bundled free" to "outcome-based experimentation", holds true as a universal business trajectory.

5.6.2 Large-Scale Quantitative Performance Linkage

A significant limitation of existing literature is the lack of empirical studies linking specific AI pricing models to audited financial performance metrics. Future research must bridge this gap by correlating the adoption of pure subscription, hybrid, or consumption-based AI models against objective financial data. Large-scale quantitative studies should investigate how these varying architectures directly impact

vital SaaS metrics, including Annual Recurring Revenue (ARR) growth, Gross Margin preservation, Net Dollar Retention (NDR), and Customer Churn. Establishing this definitive linkage will move the discourse from theoretical preference to empirically proven financial viability.

5.6.3 Experimental Pricing Studies

To further understand the behavioral dimensions of AI pricing, future scholarship should prioritize experimental pricing studies. Given the pervasive customer anxiety surrounding unpredictable "surprise bills" and the broader growth of pricing complexity, rigorous A/B testing frameworks and customer behavioral experiments are required. Researchers could empirically test how varying the presentation of dynamic pricing metrics (e.g., tokens vs. activity-based outcomes) influences user adoption and willingness to pay. Furthermore, experimental designs are ideal for investigating the ethical complexities and "black box" transparency issues associated with algorithmic pricing, testing how varying levels of price transparency affect perceived market fairness and user trust.

5.6.4 Enterprise vs. SMB AI Pricing Dynamics

This study deliberately focused on mid-sized B2B SaaS firms targeting mid-market to enterprise customers. However, the economic elasticity and purchasing behaviors of small and medium-sized businesses (SMBs) fundamentally differ from those of massive enterprise incumbents. Future research should delineate these dynamics by comparing the monetization strategies effective for SMB-focused SaaS (which typically relies heavily on Product-Led Growth and standardized tiers) against large-enterprise SaaS (which relies on bespoke, custom-negotiated consumption contracts).

5.6.5 Cost Engineering and Pricing Integration

The variable marginal costs of AI inference drive the current wave of pricing restructuring. However, the cost of LLM inference is declining rapidly, with some reports indicating year-over-year drops of over 80%. Future technical and economic research should investigate the integration of cost engineering with pricing strategy. Specifically, studies should explore how aggressive technical cost optimization such as the adoption of smaller, highly efficient open-source models versus massive proprietary LLMs, directly alters a firm's necessity to implement usage-based pricing.

5.6.6 Cross-Regional Deep Comparative Studies

Finally, while this study gathered insights from the US and Indian SaaS ecosystems, the profound heterogeneity of digital markets necessitates broader geographic inquiry. Future research should

conduct deep comparative studies exploring AI monetization strategies in highly regulated markets like Europe (navigating the implications of the GDPR and the AI Act), as well as rapidly digitizing regions such as Southeast Asia and Latin America. Understanding how varied regulatory environments, data privacy expectations, and regional software purchasing power influence the structural viability of AI pricing models remains a critical frontier for international business research.

5.7 Conclusion

The integration of Generative AI into Software as a Service is not merely a tactical feature enhancement; it represents a profound, structural evolution of the software business model. Artificial intelligence fundamentally disrupts the traditional low-marginal-cost paradigm of SaaS, replacing it with significant, variable computational expenses that mandate a complete rethink of how commercial value is captured. By empirically investigating mid-sized B2B SaaS firms, this study provides validated evidence that the industry is actively moving away from static, per-seat models toward dynamic, risk-sharing pricing architectures to accommodate this technological shift.

The empirical findings demonstrate that **hybrid pricing models** specifically the combination of baseline recurring subscriptions with metered usage or credit caps, have emerged as the dominant structural equilibrium. This hybridization is rarely driven by a desire for complex monetization, but rather acts as a necessary defensive mechanism to resolve the intense **cost-value tension** inherent in intelligent features. SaaS executives must continuously navigate profound **strategic trade-offs**, balancing the pressure to drive rapid market adoption against the critical imperative to protect gross margins from unpredictable large language model (LLM) inference costs. Furthermore, pervasive enterprise customer anxiety regarding unpredictable variable billing severely restricts the viability of pure consumption models, compelling vendors to abstract raw compute metrics into tangible, predictable business outcomes.

This research additionally establishes that AI monetization is not a static decision, but a **dynamic evolutionary pathway**. As SaaS organizations mature and their underlying infrastructure capabilities advance, they transition through sequential stages from initially bundling AI as a free, experimental growth-driver, to implementing hybrid credit allowances, and ultimately striving for sophisticated, outcome-based pricing frameworks. This trajectory underscores that capturing long-term value from active AI systems requires continuous organizational and architectural agility, moving far beyond legacy billing constraints.

Ultimately, the successful commercialization of artificial intelligence is a complex balancing act between advancing **technological capability, ensuring economic sustainability, and navigating the behavioral acceptance** of enterprise buyers. In this transformative era, pricing can no longer be treated as a downstream afterthought or a rigid, one-time operational decision. Firms that elevate monetization to a central, adaptable strategic capability investing in the necessary billing infrastructure, aligning value communication, and proactively managing variable risk, will be fundamentally better positioned to survive and thrive in the next generation of digital markets.

Chapter 6. Conclusion

6.1 Restatement of the Research Problem and Objectives

The central problem confronting the Software as a Service (SaaS) industry is a profound structural misalignment between the established economic logic of traditional software business models and the new cost dynamics introduced by embedded Artificial Intelligence (AI). Historically, SaaS economics were optimized around high fixed development costs and near-zero marginal costs of distribution, supporting predictable recurring revenue through per-seat subscriptions. However, AI-powered features introduce significant and variable marginal costs such as the computational expense of large language model (LLM) inference that scale linearly with user activity. This structural shift creates a severe tension: SaaS vendors must design pricing strategies that recover variable consumption costs to preserve their gross margins, while simultaneously accommodating enterprise customers' pervasive demands for budget predictability and resistance to unpredictable billing.

To address this challenge, the primary objective of this study was to conduct a mixed-methods empirical investigation into the AI pricing and packaging strategies adopted by mid-sized B2B SaaS companies generating between $1 million and $100 million in Annual Recurring Revenue (ARR). The research aimed to systematically map and classify prevalent monetization models to determine the specific adoption rates of pure subscription, usage-based, credit-based, and hybrid structures. Additionally, the study sought to uncover the underlying strategic, behavioral, and competitive rationales informing these executive pricing choices, culminating in the development of a practical, evidence-based decision-making framework to guide SaaS monetization.

To capture the objective market landscape alongside internal decision-making realities, the study employed a sequential mixed-methods design. The quantitative phase involved a systematic pricing-page analysis of 371 SaaS companies to establish the baseline prevalence of varying commercial architectures. Complementing this, the qualitative phase utilized in-depth, semi-structured interviews with senior SaaS executives to explore the operational trade-offs, capability constraints, and competitive pressures shaping these public-facing models.

By integrating quantitative market data with qualitative executive insights, this study sought to bridge the critical gap between pricing theory and real-world commercial practice. Ultimately, this research

provides an empirically grounded understanding of how SaaS economic models are fundamentally evolving to align with the transformative value and complex costs of intelligent systems.

6.2 Summary of Main Findings

This section distills the empirical results of the study, synthesizing the broad market realities observed in the quantitative analysis with the internal strategic rationales uncovered during the qualitative executive interviews.

6.2.1 Quantitative Findings

The systematic analysis of mid-sized B2B SaaS pricing pages established a clear, objective map of how AI features are currently commercialized. Rather than an industry-wide pivot to purely variable billing, the data reveals a heterogeneous landscape where traditional software economic models remain deeply entrenched, albeit in a state of active transition.

- **Dominant Monetization Models:** Despite industry enthusiasm for consumption-based pricing, pure subscription models remain the overwhelmingly dominant approach for monetizing AI features. Vendors predominantly embed AI capabilities into existing recurring fee structures. However, a measurable structural transition toward **hybrid monetization** is actively unfolding. The combination of a base subscription with usage-based overages or credit allocations represents the industry's emerging structural equilibrium, whereas pure, standalone usage-based models are exceedingly rare.
- **Revenue Segment Differences:** Pricing sophistication strongly correlates with organizational maturity. Firms in the early $1 million to $10 million ARR tier predominantly bundle AI for free or exhibit a high rate of “unknown” pricing, reflecting ongoing price discovery and a focus on market penetration. Conversely, as firms scale into the $50 million to $100 million ARR segment, the deployment of complex hybrid structures and credit systems scales proportionally, indicating that mature firms possess the necessary infrastructure to meter and monetize variable consumption.
- **Geographic Differences:** An evaluation of external market positioning reveals profound **global homogenization**. Public-facing pricing architectures—including packaging tiers and the dominance of US dollar pricing—are structurally uniform across different regions,

indicating that global SaaS markets have converged on a standardized positioning language regardless of whether the firm is headquartered in established or emerging digital ecosystems.

- **Packaging Patterns:** To successfully package these capabilities, tiered pricing remains the standard, utilizing AI as a distinct differentiation lever to drive customers into higher-tier enterprise plans. Furthermore, the data shows a deliberate avoidance of highly abstract, technical metrics (like AI tokens); instead, firms rely heavily on **credit systems** or user seats, utilizing them as an abstraction layer to translate complex computational costs into comprehensible purchasing units.

6.2.2 Qualitative Findings

While the quantitative data maps the structural reality of the market, the qualitative interviews with SaaS executives explain the strategic tensions and behavioral constraints driving these specific choices. The thematic analysis surfaced several core operational realities:

1. **Margin Protection vs. Adoption Trade-off:** Executives are acutely aware of the unpredictable, variable infrastructure costs associated with large language model (LLM) inference. They face a persistent tension between aggressively driving AI adoption and shielding their gross margins from power users. To manage this, leaders frequently deploy internal guardrails, temporarily absorbing costs for basic features to build user habits while strictly metering highly complex, deterministic AI tasks to prevent unrecoverable financial losses.
2. **Customer Price Predictability Concerns:** A major barrier to dynamic monetization is profound customer resistance to unpredictable billing. Enterprise buyers actively reject the financial risk of uncapped usage models and struggle to map granular compute metrics to business value. Executives repeatedly emphasized that commercial simplicity and budget predictability are prerequisites for B2B adoption, explaining the heavy reliance on flat-rate subscriptions or distinct, outcome-aligned "jobs to be done" rather than raw AI usage.
3. **Competitive Anchoring:** The pricing power of B2B SaaS firms is heavily restricted by external market pressures. The ubiquity of low-cost, consumer-grade GenAI tools has anchored enterprise willingness to pay at artificially low rates. Consequently, executives are frequently

forced to offer foundational AI capabilities at no additional charge simply to maintain competitive feature parity and avoid customer churn, effectively treating basic AI as a baseline expectation rather than a premium differentiator.

Ultimately, synthesizing these findings demonstrates that the quantitative patterns represent the *what*, a heavy reliance on hybrid and subscription models, while the qualitative themes explain the *why*: these models act as pragmatic, defensive compromises engineered to satisfy the vendor's need for cost recovery while catering to the enterprise buyer's demand for absolute predictability.

6.3 Implications of the Research

This section translates the empirical findings into actionable insights, detailing how the integration of embedded AI features structurally impacts academic frameworks of software economics and providing practical guidance for SaaS practitioners.

6.3.1 Theoretical Implications

This study formally advances the academic understanding of software economics by explicitly extending SaaS pricing theory into variable-cost environments. Historically, literature on software pricing and cost estimation, such as the algorithmic models validated by Kemerer (1986), operated on the foundational assumption of high upfront fixed development costs and near-zero marginal distribution expenses. The empirical results of this research confirm that the embedding of active AI features introduces significant, variable marginal costs that scale linearly with usage, fundamentally invalidating these legacy assumptions. By documenting how firms adapt their packaging to this structural break, the findings extend prior conceptualizations of dynamic software pricing presented by Esan (2021), demonstrating that AI integration forces a necessary departure from static value capture models.

Furthermore, this research provides vital empirical grounding to the literature on SaaS pricing heterogeneity and complexity discussed by Saltan & Smolander (2021) and García-Fernández et al. (2025). The study quantifies the structural shift toward risk-sharing mechanisms, explicitly validating hybrid monetization as a pragmatic, transitional equilibrium that resolves the tension between capturing variable value and satisfying enterprise demands for predictability. While outcome-based pricing remains the theoretical ideal, this research confirms it is currently limited in practice due to attribution complexities, further cementing the hybrid model's role as the dominant risk-mitigation tool. Finally, by documenting the evolutionary transition from bundled loss-leaders to complex outcome-based metrics, the study enriches the AI innovation typology and business model transformation theories established by Paschen et al. (2020) and Mishra & Tripathi (2021), grounding abstract conceptual claims in observable commercial realities.

6.3.2 Managerial Implications

The practical challenge of monetizing intelligent features has rapidly elevated from a routine product-level decision to a critical board-level imperative for SaaS leadership (Metronome, 2025). To assist executives in navigating this transition, the findings yield several highly practical takeaways regarding strategic decision-making, margin management, and customer communication.

Strategic Decision Framework and Stage-Based Recommendations: Monetization strategies must be dynamically aligned with firm maturity and customer predictability demands. For **early-stage SaaS firms** ($1M–$10M ARR), leaders should initially bundle foundational AI features into core subscriptions to drive immediate adoption, establish user habits, and train models, temporarily absorbing the compute costs. As firms scale into the **mid-market**, they must transition toward premium tier gating or isolating advanced capabilities as distinct add-ons to begin recovering costs. For **mature SaaS firms** nearing the $100 million ARR threshold, the deployment of sophisticated hybrid architectures—specifically those utilizing usage-based overages or credit allowances—becomes an absolute necessity to capture value from power users without alienating the broader customer base.

Margin Management Guidance: SaaS leaders must proactively deploy structural safeguards to avoid catastrophic "margin traps" associated with unpredictable compute consumption (McKinsey & Company, 2025; Bain & Company, 2025). Executives are advised to implement strict internal cost modeling and design clear overage policies, such as providing a fixed, pooled allowance of AI credits within a baseline subscription. The utilization of these credit systems serves as a vital abstraction layer; by translating complex, variable token costs into predictable, psychological pricing units, firms can protect their gross margins while maintaining pricing formats that enterprise buyers easily comprehend.

Customer Communication: Finally, go-to-market teams must deliberately shift away from selling raw technological metrics. To overcome enterprise anxiety, customer communication must heavily emphasize framing the Return on Investment (ROI) and positioning AI capabilities as tangible productivity multipliers. Vendors must ensure that AI pricing maps directly to specific business outcomes or "jobs to be done" to avoid the perception of imposing an unpredictable "AI tax" on existing workflows. Ultimately, firms that elevate their billing infrastructure and align their value communication will be fundamentally better equipped to survive this economic transition.

6.4 Limitations of the Study

While this research provides robust, empirically grounded insights into how mid-sized B2B SaaS firms are monetizing embedded Artificial Intelligence, it is subject to several methodological and contextual constraints. These limitations are outlined transparently to ensure academic rigor and define the precise boundaries of the study's applicability.

First, the quantitative phase is inherently constrained by its reliance on public-facing data. By systematically analyzing the pricing pages of 371 companies, the study introduces a "public pricing page bias". This methodology successfully captures external market positioning but inevitably obscures the reality of custom-negotiated enterprise contracts, off-book discounting, and the complex structures gated behind "contact sales" barriers. Additionally, while the sample was deliberately scoped to firms generating between $1 million and $100 million in Annual Recurring Revenue (ARR), the reliance on external revenue estimates means these categorizations may not be perfectly precise.

Second, the qualitative phase drew upon a limited sample size of senior SaaS executives. While these in-depth interviews were essential for uncovering the strategic rationales and operational tensions driving pricing decisions, this small sample size restricts the broad generalizability of the internal practices discussed across the entire SaaS ecosystem and introduces a natural risk of response bias. Furthermore, these qualitative insights depend entirely on self-reported performance metrics; executive claims regarding margin control or customer adoption carry a potential for perception bias or inflation, as the study lacked access to audited financial data or internal product telemetry.

Third, this research employs a cross-sectional design, capturing a specific "snapshot in time" of a highly fluid market. Because the underlying costs of LLM inference are declining rapidly and technological capabilities are in constant flux, the specific pricing archetypes and thresholds documented here are subject to rapid, ongoing evolution.

Finally, the study's geographic scope focused primarily on the United States and India. Consequently, the strategic nuances observed may not perfectly generalize to other international markets governed by different regulatory environments, data privacy laws, or software purchasing behaviors. Ultimately, these limitations do not invalidate the study's core findings; rather, they transparently define the boundaries of inference and contextualize the proposed decision-making framework.

6.5 Recommendations for Future Research

The findings and limitations of this study expose critical gaps in the academic understanding of AI monetization, offering several high-value avenues for future inquiry. As technological capabilities and cost structures evolve rapidly, subsequent research must expand beyond cross-sectional observations to capture the dynamic nature of this commercial transition.

- **Longitudinal AI Pricing Evolution Studies:** Given the unprecedented technological evolution and cost volatility of underlying large language models, current cross-sectional data represents only a snapshot in time. Future research should employ longitudinal methodologies to track specific SaaS cohorts over a 3-to-5-year period. This would definitively capture how firms dynamically adjust their pricing architectures as they move from early experimentation toward structural maturity.
- **Quantitative Performance Linkage:** A significant limitation of existing literature is the lack of empirical studies linking specific AI pricing models to audited financial performance. Future large-scale quantitative studies must bridge this gap by using regression analysis to correlate the adoption of pure subscription, hybrid, or consumption-based models against vital SaaS metrics, including Annual Recurring Revenue (ARR) growth, Gross Margin preservation, Net Dollar Retention (NDR), and Customer Churn.
- **Enterprise SaaS Segment Analysis:** This study deliberately focused on mid-sized B2B SaaS firms ($1M–$100M ARR) relying heavily on public pricing pages. Future research should investigate massive enterprise incumbents, exploring how custom-negotiated consumption contracts and opaque monetization structures are deployed at the highest end of the market.
- **Sector-Specific Studies:** Future research should delineate monetization dynamics by comparing Horizontal SaaS platforms with industry-specific Vertical SaaS solutions. As noted in the qualitative findings, vertical solutions often successfully map pricing to specific user roles or "jobs to be done", suggesting that vertical software may utilize fundamentally different monetization frameworks than broad, horizontal tools.
- **Customer-Side Research:** To fully understand the behavioral dimensions of algorithmic pricing, future scholarship should prioritize customer-side behavioral research. Researchers should empirically investigate enterprise buyer perceptions of AI pricing fairness,

particularly addressing the transparency and ethical complexities inherent in "black box" usage models.

- **Deeper Geographic Comparisons:** Finally, the profound heterogeneity of digital markets necessitates broader geographic inquiry. Future comparative studies should contrast AI monetization strategies in highly regulated markets, such as Europe, with rapidly digitizing emerging regions to understand how varying regulatory environments, data privacy expectations, and purchasing power influence commercial viability.

6.6 Final Thoughts and Reflections

The rapid integration of Generative AI into Software as a Service fundamentally alters the economic architecture of the industry, representing a profound structural shift rather than a mere tactical feature enhancement. By introducing significant and variable computational costs, AI effectively dismantles the traditional low-marginal-cost paradigm that has governed software economics for decades.

As this research demonstrates, the SaaS industry remains in a highly transitional phase of active experimentation. There is no single, universally dominant monetization model that perfectly resolves the complexities of AI costs. Instead, firms are navigating a fluid landscape, continuously iterating their commercial structures as foundational models become more efficient and enterprise buyers become more accustomed to algorithmic tools.

The central insight of this study is that **hybrid monetization**—specifically the blending of stable recurring subscriptions with capped usage or credit allocations—has emerged as a **vital, pragmatic equilibrium**. This hybridization serves as a necessary defensive mechanism, bridging the critical gap between a vendor’s need to recover variable infrastructure costs and an enterprise customer's steadfast demand for budget predictability. Ultimately, the firms succeeding in this environment are those that look beyond raw compute metrics, strategically aligning their dynamic cost structures directly with tangible, perceived customer value.

In conclusion, the challenge of monetizing artificial intelligence is **not merely a pricing tweak, but a complete redesign of the SaaS business model**. As software transitions from a passive human-operated tool to an active, intelligent collaborator, the mechanisms for commercializing it must evolve accordingly. For the next generation of digital markets, the ultimate measure of strategic success will

not just be technological innovation, but the organizational agility to continuously align value creation with sustainable value capture in intelligent systems.

Definitions of Key Terms

Artificial Intelligence (AI): Technologies that use machine learning, natural language processing, or related methods to automate or augment cognitive tasks.

Generative AI: AI systems capable of producing new content such as text, code, images, or predictions, typically based on large language models.

Software as a Service (SaaS): A software delivery model in which applications are hosted in the cloud and accessed by users through subscription-based pricing.

AI Feature: Any functionality within a SaaS product that is powered by AI or machine learning.

Pricing Model: The structure through which a company charges customers for software usage or access.

Usage-Based Pricing: A variable pricing model in which customers pay based on consumption metrics such as tokens, API calls, predictions, or tasks completed.

Outcome-Based Pricing: A pricing model that links charges to measurable business results achieved through AI capabilities.

Bibliography

Bain & Company. (2025). *Per-seat software pricing isn't dead, but new models are gaining steam.* Bain & Company. https://www.bain.com/insights/per-seat-software-pricing-isnt-dead-but-new-models-are-gaining-steam/

BCG. (2025). *The evolution of pricing models for SaaS companies.* BCG. https://medium.com/bcgontech/the-evolution-of-pricing-models-for-saas-companies-6d017101d733

Bessemer Venture Partners. (2023). *The rise of SaaS in India 2023.* Bessemer Venture Partners. https://www.bvp.com/atlas/rise-of-saas-in-india-2023

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

Coherent Solutions. (2025). *AI development cost estimation: Pricing structure, implementation ROI.* Coherent Solutions. https://www.coherentsolutions.com/insights/ai-development-cost-estimation-pricing-structure-roi

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Esan, O. (2021). Dynamic pricing models in SaaS: A comparative analysis of AI-powered monetisation strategies. *International Journal of Research Publication and Reviews, 2*(12), 1757–1772.

Fortune Business Insights. (2025). *Software as a service (SaaS) market size, global report, 2032.* Fortune Business Insights. https://www.fortunebusinessinsights.com/software-as-a-service-saas-market-102222

García-Fernández, A., Parejo, J. A., Trinidad, P., & Ruiz-Cortés, A. (2025). Automated analysis of pricings in SaaS-based information systems. *arXiv.* https://doi.org/10.48550/arXiv.2503.21444

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

IJCESEN. (2025). The ethics of AI in pricing: Fairness, transparency, and accountability. *International Journal of Computational and Experimental Science and Engineering.* https://www.ijcesen.com/index.php/ijcesen/article/view/3949

Kemerer, C. F. (1987). An empirical validation of software cost estimation models. *Communications of the ACM, 30*(5), 416–429. https://doi.org/10.1145/22899.22906

Khanna, M., Mittal, N., Wu, A., & Castillo, M. (2025). *Upgrading software business models to thrive in the AI era.* McKinsey & Company. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/upgrading-software-business-models-to-thrive-in-the-ai-era

L.E.K. Consulting. (2025). *The future role of generative AI in SaaS pricing.* L.E.K. Consulting. https://www.lek.com/insights/tmt/us/ei/future-role-generative-ai-saas-pricing

McKinsey & Company. (2025). *Evolving models and monetization strategies in the new AI SaaS era.* McKinsey & Company. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/upgrading-software-business-models-to-thrive-in-the-ai-era

Metronome. (2025). *AI pricing in practice: 2025 field report from leading SaaS teams.* Metronome. https://metronome.com/blog/ai-pricing-in-practice-2025-field-report-from-leading-saas-teams

Mishra, S., & Tripathi, A. R. (2021). AI business model: An integrative business approach. *Journal of Innovation and Entrepreneurship, 10*(1), 18. https://doi.org/10.1186/s13731-021-00157-5

Paschen, J., Wilson, M., & Ferreira, J. J. (2020). Collaborative intelligence: How human and artificial intelligence create value along the B2B sales funnel. *Business Horizons, 63*(3), 403–414. https://doi.org/10.1016/j.bushor.2019.12.003

Paschen, U., Pitt, C., & Kietzmann, J. (2020). Artificial intelligence: Building blocks and an innovation typology. *Business Horizons, 63*(2), 147–155. https://doi.org/10.1016/j.bushor.2019.10.004

SaaStr / Unusual Ventures. (2025). *How to price and package AI SaaS products.* SaaStr. https://www.saastr.com/how-to-price-and-package-ai-saas-products/

Saltan, A., & Smolander, K. (2021). How SaaS companies price their products: Insights from an industry study. In E. Klotins & K. Wnuk (Eds.), *Software Business: 11th International Conference, ICSOB 2020* (pp. 1–13). Springer. https://doi.org/10.1007/978-3-030-67292-8_1

Valueships. (2025). *The best pricing for SaaS: Models, strategies and tips for success.* Valueships. https://www.valueships.com/post/ai-saas-pricing-guide-learn-from-openais-mistake

Zuora. (2025). *Monetizing AI: Why SaaS companies are missing revenue opportunities.* Zuora. https://www.zuora.com/resource/monetizing-gen-ai-why-saas-companies-are-missing-out/

Appendix A: Quantitative Data

Company Distribution By Revenue

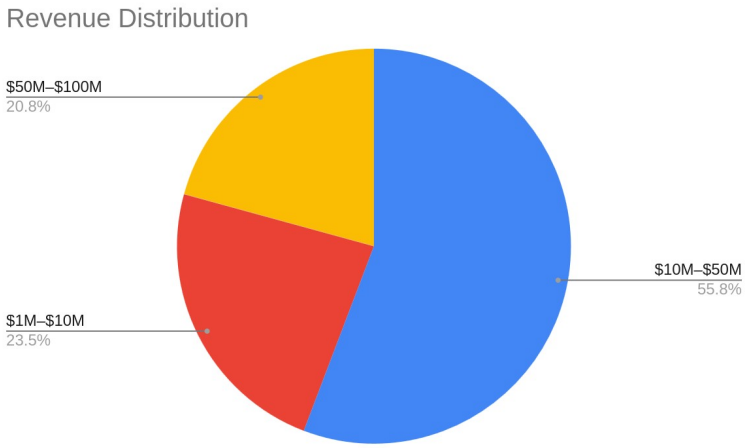


Figure 7: Company Distribution By Revenue

Table 4: Company Distribution By Revenue

Revenue Range	# of companies
$10M–$50M	207
$1M–$10M	87
$50M–$100M	77
Grand Total	**371**

Pricing Strategy Distribution

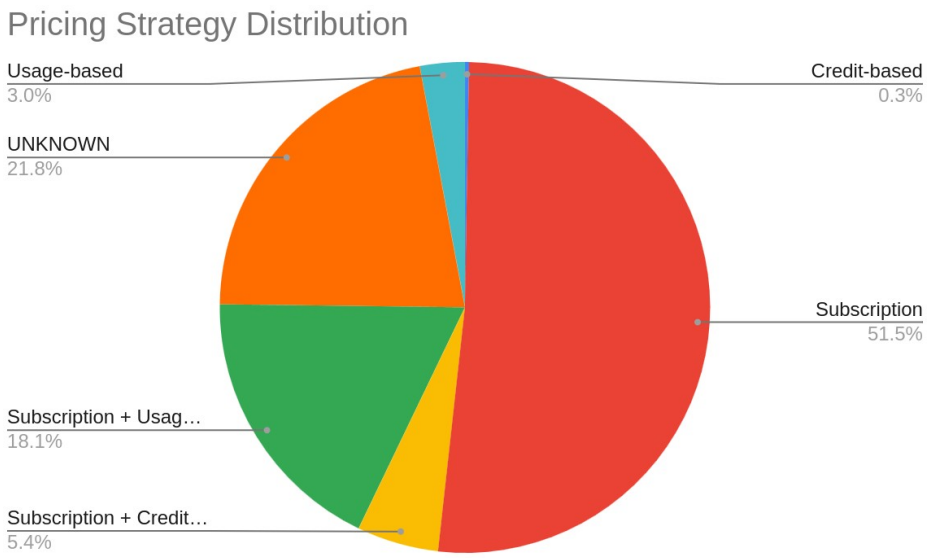


Figure 8: Pricing Strategy Distribution

Table 5: Pricing Strategy Distribution

Pricing Strategy	# of Companies
Credit-based	1
Subscription	191
Subscription + Credit-based	20
Subscription + Usage-based	67
UNKNOWN	81
Usage-based	11
Grand Total	**371**

Pricing Strategy by Revenue Range

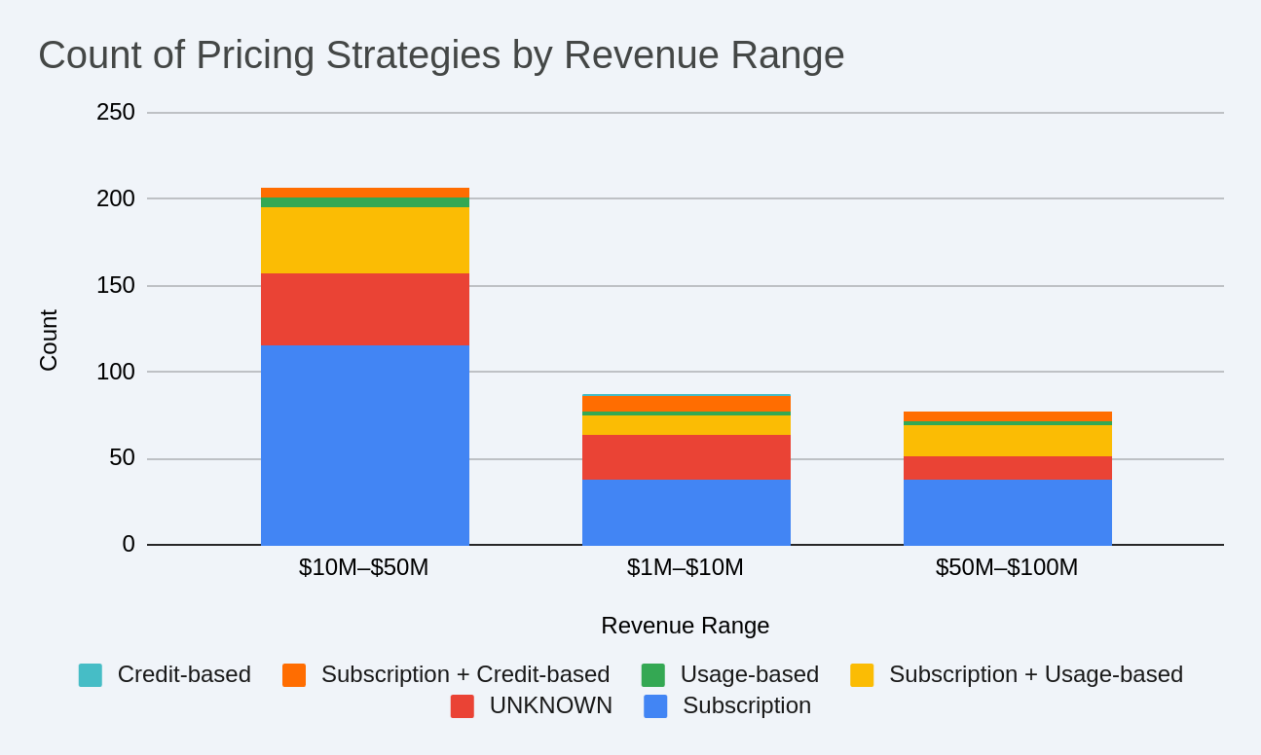


Figure 9: Pricing Strategy by Revenue Range

Table 6: Pricing Strategy by Revenue Range

Revenue Range	Credit-based	Subscription	Subscription + Credit-based	Subscription + Usage-based	UNKNOWN	Usage-based	Grand Total
$10M–$50M		115	6	38	42	6	207
$1M–$10M	1	38	9	11	26	2	87
$50M–$100M		38	5	18	13	3	77
Grand Total	**1**	**191**	**20**	**67**	**81**	**11**	**371**

Features Provided by Reviewed Companies

Table 7: Features Provided by Companies

Features	COUNT of Features
Adaptive Alerts	1
AI Assistant	59
AI Chatbot	1
AI coverage summaries	1
AI Image Recognition	1
AI press release reviews	1
AI-powered detection	1
Audio Generation	1
Automation	51
Behavioral highlights	1
Content Editing	21
Conversational AI	1
Customer Segmentation	1
Generative AI	1
Image Generation	11
Intelligent error grouping	1
Natural Language Processing	1
Other Generative AI	9
Other Machine Learning	18
Predictive Analytics	43
Recommendations	16
Search and Retrieval	16
Speech to Text	4
Suspicious activity detection	1
Text Generation	45
Text to Speech	5
Transcription	9
Video Generation	8
Work time classification	1

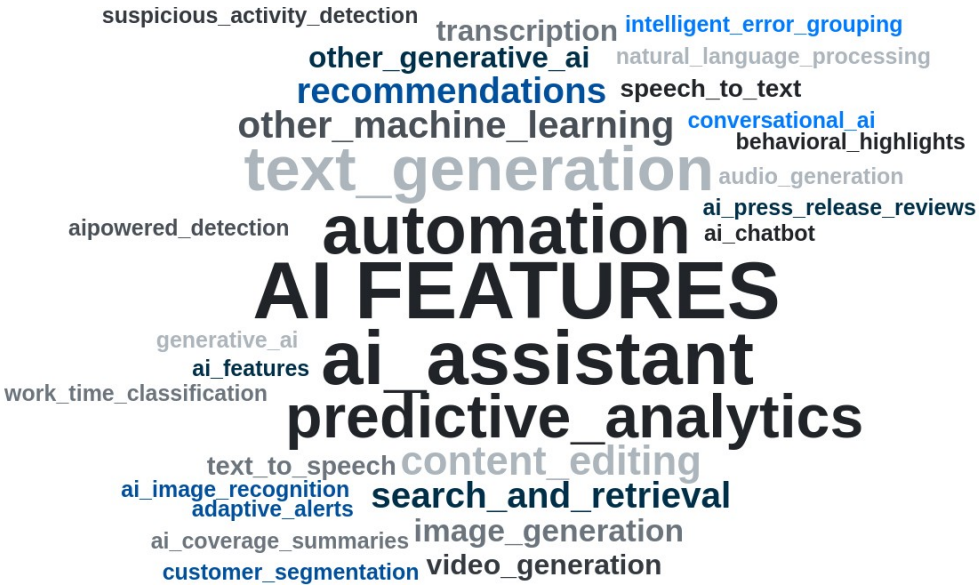


Figure 10: Features

Provided by Reviewed Companies

Appendix B: Qualitative Assessment

Questionaire

Monetizing AI in SaaS

This questionnaire is part of a doctoral research study on how B2B SaaS companies design, price, and package AI-enabled features. All responses are confidential and for academic use only.

1. Optional contact email

If you would like to receive a summary of the research findings upon completion of the study, you may optionally provide an email address below. This information will be collected separately and will not be linked to your questionnaire responses in any way.

Privacy & Confidentiality

To encourage candid and unbiased responses from senior executives, this study intentionally minimizes the collection of personally identifiable information. Contact details are not required for participation.

All data are used exclusively for academic research purposes and handled in accordance with institutional ethical guidelines.

By proceeding with this questionnaire, you indicate your informed consent to participate in this study.

SECTION 1: COMPANY CONTEXT

Q1. Please briefly describe your company's primary product and target customer segment.

Q2. Have you shipped any AI-powered features or enhancements in your core SaaS product?

Mark only one oval.

- ● Yes *Skip to question 4*
- ● No *Skip to question 12*

SECTION 2: AI MONETIZATION

Q3. Do you charge customers extra for AI-powered features? *

Mark only one oval.

- ● Yes *Skip to question 6*
- ● No *Skip to question 5*
- ● Partially (some AI features are charged; others are bundled) *Skip to question 6*

SECTION 2A: BUNDLED AI RATIONALE

Q3a. What influenced your decision to include AI features without additional charges? *

SECTION 3: PRICING STRATEGY AND TRADE-OFFS

Q4. How did you decide whether AI should be bundled, sold as an add-on, or priced using usage- or credit-based models?

Q5. What have been the most significant trade-offs or considerations in pricing and packaging AI features? (e.g., infrastructure cost, customer willingness to pay, competitive pressure, pricing predictability)

Q6. What outcomes or impacts have you observed from your current AI monetization approach? (e.g., adoption, margins, sales cycles, churn, positioning)

SECTION 4: MARKET AND COMPETITIVE CONTEXT

Q7. How have competitors or broader market expectations influenced your AI pricing or packaging decisions?

Q8. Have you observed regional differences in willingness to adopt or pay for AI features? If yes, please explain briefly.

SECTION 5: FUTURE OUTLOOK

Q9. How do you expect your AI pricing and packaging strategy to evolve over the next 12–24 months?

SECTION 6: FINAL REFLECTION (ALL RESPONDENTS)

Q10. Based on your experience, what advice would you offer to SaaS leaders who are deciding how to monetize AI features?
