IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 19 June 2025, accepted 6 July 2025, date of publication 10 July 2025, date of current version 16 July 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3587943

RESEARCH ARTICLE

PRISE: A Framework for AI Product Incubation From Concept to Implementation

BAOLI WANG, YE LI, (Member, IEEE), AND WEN LEI
Microsoft Search Technology Center Asia, Haidian, Beijing 100080, China
Corresponding author: Ye Li (jull@microsoft.com)

ABSTRACT AI product incubation faces unique challenges such as difficulty in technical feasibility assessment, ambiguous value perception, and engineering transformation barriers. This research proposes the PRISE framework (Problem Definition, Rapid Validation, Impact Evaluation, Single-focus Development, End-to-End Productization), integrating design thinking, lean startup, and engineering practices with adaptive modifications for AI characteristics. Through analysis of seven cross-industry cases from Microsoft's AI Vertical team, we validated the framework's practicality. The research reveals five key patterns for successful AI product incubation: user-decision maker separation identification, lightweight validation priority, technology-value transformation mechanism, deep cultivation of vertical scenarios, and end-to-end delivery thinking. This framework fills the gap in AI product development methodology, providing systematic guidance and operational tools for practitioners while offering new perspectives for AI innovation management theory.

INDEX TERMS Artificial intelligence, design thinking, innovation, lean startup, product development.

I. INTRODUCTION

Artificial intelligence technology has made significant advancements in recent years, yet the practical implementation of AI in enterprises faces major challenges. According to industry research, over 80% of AI projects fail—twice the failure rate of non-AI IT projects [10]. Gartner surveys reveal that on average only 48% of AI projects reach production [7]. By the end of 2025, at least 30% of generative AI projects will be abandoned after proof-of-concept due to poor data quality, inadequate risk controls, escalating costs, or unclear business value [6]. The Global CDO Insights 2025 survey identifies the main obstacles as insufficient data quality and readiness (43%), inadequate technical maturity (43%), and lack of skills and data literacy (35%) [10]. These findings suggest that AI project failures stem primarily from a lack of systematic incubation frameworks rather than the technology itself.

Existing product development methodologies such as design thinking [3] and lean startup [5] demonstrate obvious limitations when applied to AI products: design thinking neglects AI systems' high dependency on data distribution and computational resources; lean startup does not adequately consider the longer training and optimization cycles characteristic of AI models. This misalignment between methodologies and technological characteristics exacerbates the difficulties in transitioning AI products from concept to implementation.

Addressing these challenges, this research focuses on three core questions: (1) What are the unique challenges and barriers in the AI product incubation process? (2) How can we construct a systematic incubation methodology adapted to AI product characteristics? (3) How can this methodology be applied and validated in actual projects? Accordingly, our research objectives are to identify key challenges in AI product transition from concept to implementation, construct an incubation framework integrating design thinking, lean startup, and engineering practices, validate its applicability through actual cases, and provide operational methodological tools for AI product development teams.

This research adopts a multiple-case study method [25], systematically analyzing seven AI product incubation projects by Microsoft's AI team across education, healthcare, advertising, retail, and other industries. Data was collected

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei.

through in-depth interviews, document analysis, and participatory observation, combined with theory-oriented thematic analysis to construct the PRISE framework. The framework's effectiveness was examined through cross-case retrospective validation, key indicator verification, and external review validation mechanisms.

The research finds that AI product incubation faces unique challenges including difficulties in technology and business integration, weak user value perception, and missing productization transformation mechanisms. The PRISE framework effectively addresses these challenges through five key dimensions (problem definition, rapid validation, impact evaluation, single-focus development, and end-to-end productization). Successful AI product incubation requires the triple integration of front-loaded technical feasibility assessment, value verification closed loops, and engineering genes, providing a systematic path applicable to AI product development across different industries.

The structure of this paper is as follows: Chapter 2 reviews the theoretical foundations and explains the PRISE framework construction logic; Chapter 3 describes the research methodology and case selection criteria; Chapter 4 demonstrates the framework's practical application through seven cases; Chapter 5 explores the framework's application mechanisms and validation results; Chapter 6 discusses research limitations and future directions.

II. THEORETICAL FOUNDATION AND FRAMEWORK CONSTRUCTION

A. EXISTING APPROACHES AND THEIR LIMITATIONS IN AI PRODUCT DEVELOPMENT

1) TRADITIONAL PRODUCT DEVELOPMENT METHODOLOGIES

AI product development involves three key theoretical perspectives, each revealing unique challenges. From a product management perspective, traditional theories [8], [11] face challenges in the AI field including more complex balance between user needs and technical feasibility, iteration cycles constrained by model training, and difficulty in establishing user trust [19]. From a software engineering perspective, AI system development presents three major differences: strong data dependency, difficult testing validation, and high explainability requirements [1], while also needing to address ''hidden technical debt'' issues [23]. From an AI system characteristics perspective, data quality, model explainability, engineering deployment difficulty, value assessment complexity, and ethical compliance constitute five unique challenges [14], with the core dilemma being bridging the gap between technological possibilities and user expectations [13].

Existing product development methodologies exhibit significant limitations when applied to AI. Design thinking [3], while advantageous in user needs discovery, neglects technical feasibility assessment in AI product development, cannot support rapid prototype iteration, and struggles to manage user expectations. Lean startup [5] faces three major challenges in the AI domain: ignoring data dependency, feedback cycle mismatches, and disconnection between technical metrics and user value. Agile development also encounters obstacles in AI project implementation due to highly interdisciplinary teams, high technical uncertainty, ambiguous delivery definitions, and complex continuous integration [31].

2) AI-SPECIFIC LIFECYCLE MANAGEMENT APPROACHES

Recent developments in AI lifecycle management have introduced specialized approaches that address some technical aspects of AI system development. MLOps (Machine Learning Operations) focuses on the operationalization of machine learning models through automated pipelines, continuous integration/continuous deployment (CI/CD) for ML, model monitoring, and version control [12]. MLOps provides robust solutions for model deployment, monitoring, and maintenance but primarily addresses the technical infrastructure aspects of AI systems post-development.

Data-Centric AI represents an emerging paradigm that emphasizes the systematic improvement of data quality rather than focusing solely on model architecture optimization [17], [30]. This approach advocates for structured data management, systematic labeling strategies, and iterative data quality improvement [29]. Research has shown that data quality issues create ''data cascades'' that compound throughout AI system development [22]. Data-Centric AI provides valuable insights into the critical role of data in AI system performance but primarily focuses on the technical dimensions of data engineering.

Table 1 presents a systematic comparison of these AI lifecycle approaches with the PRISE framework, highlighting their complementary nature and distinct focus areas.

The comparison reveals that while MLOps and Data-Centric AI provide essential technical foundations, they address different challenges than PRISE. MLOps excels in technical infrastructure and operational excellence but lacks user-centered design and business value validation mechanisms. Data-Centric AI provides systematic data management approaches but does not address user needs discovery, market validation, or product strategy considerations. PRISE integrates these technical considerations within a broader product development context, providing the missing bridge between technical capabilities and business value realization.

These mismatches between existing methodologies and AI product characteristics indicate the need for a systematic framework that integrates the advantages of existing theories while specifically adapting to the unique challenges of AI.

B. PRISE FRAMEWORK CONSTRUCTION AND ANALYSIS

The PRISE framework addresses gaps across multiple levels of AI product development methodologies. While AI-specific technical approaches (MLOps, Data-Centric AI) provide operational foundations, and traditional product

TABLE 1. Comparison of PRISE with AI lifecycle management approaches.

Comparison Dimension	MLOps	Data-Centric AI	PRISE Framework
Primary Focus	Model deployment & operations	Data quality optimization	End-to-end product incubation
Lifecycle Coverage	Post-development operations	Data preparation & engineering	Concept to production
Core Activities	CI/CD for ML, monitoring, versioning	Data labeling, quality assessment, iterative data improvement	Problem definition, validation, value assessment, productization
Business Integration	Limited business context	Minimal business considerations	Strong business-technical integration
User Perspective	Engineering-centric	Technical team-centric	User and stakeholder-centric
Validation Approach	Technical performance metrics	Data quality metrics	Multi-level value validation
PRISE Complementarity	Supports End-to-End Productization dimension	Supports Problem Definition technical anchoring	Provides holistic framework

methodologies (Design Thinking, Lean Startup) offer innovation processes, neither adequately addresses the unique intersection of technical feasibility, user value, and business viability in AI product development. PRISE bridges these gaps through a systematic integration approach.

1) FRAMEWORK OVERVIEW

The PRISE framework integrates the advantages of design thinking, lean startup, and software engineering methodologies with adaptive modifications for AI product characteristics. As illustrated in Figure 1, PRISE encompasses five core dimensions: Problem Definition, Rapid Validation, Impact Evaluation, Single-focus Development, and End-to-End Productization. These five dimensions constitute both the stages of AI project incubation and correspond to five core product design questions: what to build, how to test, is it valuable, where to start, and how to deliver.

The framework adopts a ''three-phase integration'' strategy: during the needs exploration period, it incorporates design thinking empathy methods with front-loaded technical feasibility assessment; during the experimental validation period, it inherits the hypothesis-driven approach from lean startup while strengthening value verification closed loops; during the engineering delivery period, it adopts agile practices while emphasizing AI system-specific engineering requirements.

2) CORE MECHANISMS OF THE FIVE DIMENSIONS

Problem Definition stage focuses on identifying real users and discovering pain points. Key activities include user role and decision chain analysis, pain point hierarchical analysis, solution hypothesis construction, and front-loaded technical feasibility assessment. PRISE particularly emphasizes technical anchoring mechanisms, prioritizing data availability assessment to avoid designing technically infeasible solutions.

Rapid Validation stage employs low-cost methods to verify hypotheses. Key activities include minimum viable hypothesis design, lightweight experiment design, user feedback collection, and hypothesis iterative adjustment. PRISE focuses on AI-specific validation approaches, such as using pre-trained models or Wizard of Oz methods to quickly build functional demonstrations, avoiding the high costs of complete AI system development.

Impact Evaluation stage concentrates on product usability and perceived value. Core activities include establishing value assessment frameworks, building metric conversion mechanisms, testing user value perception, and designing continuous evaluation mechanisms. PRISE emphasizes building conversion models from technical metrics → behavioral metrics → business metrics, ensuring technology capability improvements translate into business value.

Single-focus Development stage emphasizes focusing on a single vertical scenario in the initial phase. Key activities include vertical scenario selection, core functionality definition, end-to-end process construction, and expansion path planning. By focusing on small but practical entry points, teams can quickly achieve closed-loop validation and form successful samples.

End-to-End Productization stage integrates algorithmic, engineering, and product capabilities, focusing on transitioning from concept to production-ready systems. Key activities include engineering architecture design, quality assurance mechanism establishment, deployment and operations planning, and user feedback loop construction. PRISE particularly emphasizes engineering genes, requiring consideration of monitoring, rollback, and other production-ready capabilities from the initial product stages.

3) KEY TERMINOLOGY AND CONCEPTUAL CLARIFICATIONS

To ensure clarity and consistency throughout this framework, we define the following key terms that are central to the PRISE methodology:

Technical Anchoring Mechanism: A systematic process for evaluating technical feasibility constraints (particularly data availability, computational requirements, and

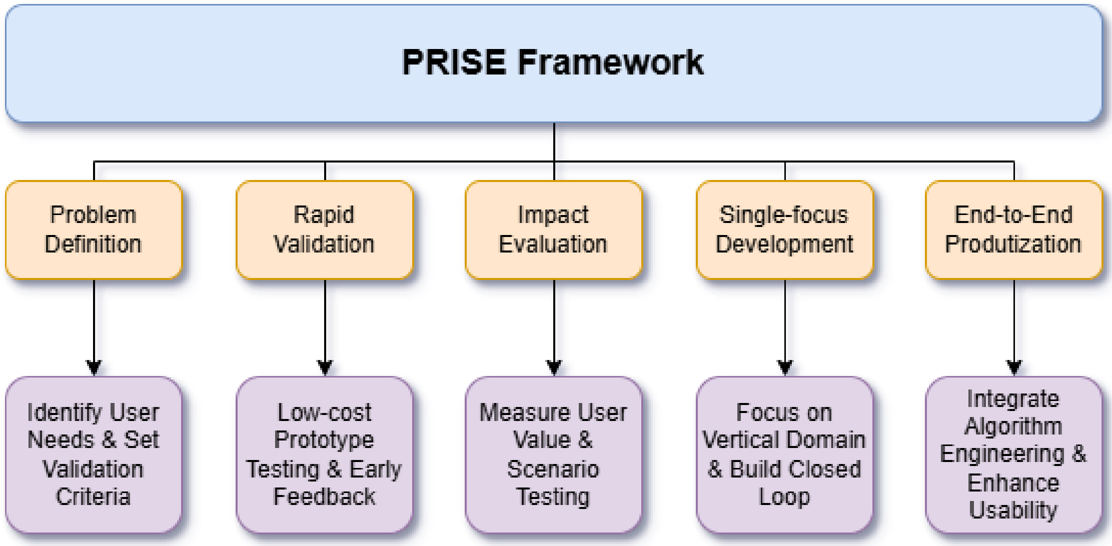


FIGURE 1. PRISE framework: A systematic approach for AI product incubation.

algorithmic maturity) during the early problem definition stage. This mechanism prevents teams from pursuing innovation directions that are technically infeasible or require resources beyond organizational capacity. It differs from traditional feasibility studies by being AI-specific, focusing on data quality, model performance expectations, and deployment complexity.

Value Verification Closed Loop: A three-tier validation system that establishes clear conversion pathways from technical metrics (e.g., model accuracy, response time) to behavioral metrics (e.g., user engagement, task completion) to business metrics (e.g., revenue impact, cost reduction). This closed loop ensures that AI capability improvements translate into measurable business outcomes rather than remaining isolated technical achievements.

Engineering Genes: Production-ready capabilities that are embedded into AI products from the design stage rather than added retroactively. These include monitoring systems, rollback mechanisms, error handling, scalability considerations, and continuous integration practices. The term "genes" emphasizes that these capabilities should be inherent to the product's architecture, not external additions.

Lightweight Validation: Low-cost, rapid experimentation methods specifically adapted for AI product validation, including Wizard of Oz prototyping (simulating AI capabilities through human operators), pre-trained model demonstrations, and synthetic data testing. These approaches enable hypothesis testing without full AI system development.

Hypothesis-Driven Resource Optimization: A decision-making approach that requires explicit hypothesis formulation before resource allocation, clear success criteria definition, and systematic validation before scaling investment. This prevents the common AI development pitfall of over-investing in technically interesting but commercially unviable directions.

Stakeholder Role Differentiation: The systematic identification and analysis of different stakeholder types in AI product ecosystems, particularly distinguishing between end users (who interact with the system), decision makers (who approve adoption), technical evaluators (who assess implementation feasibility), and value recipients (who benefit from outcomes). This differentiation is critical because AI products often involve complex organizational decision chains.

These terms represent core conceptual innovations within the PRISE framework and should be understood as interconnected components of a systematic approach to AI product development rather than isolated techniques.

4) DIFFERENTIATING CHARACTERISTICS OF THE PRISE FRAMEWORK

The PRISE framework has three major differentiating characteristics compared to existing methodologies, as shown in Table 2. While Table 1 demonstrated PRISE's complementary relationship with AI-specific technical approaches, this comparison focuses on fundamental product development methodologies to highlight PRISE's core innovations. The PRISE framework's three innovative characteristics include:

- **Technical Anchoring Mechanism**: Introduces data availability assessment during the problem definition stage to avoid "technically infeasible innovation." For example, in educational product design, teams are required to evaluate training data availability and quality before determining product direction, preventing project failure due to insufficient data later.
- **Value Verification Closed Loop**: Establishes a conversion model from technical metrics → behavioral

TABLE 2. Comparison of PRISE with major methodologies.

Comparison Dimension	Design Thinking	Lean Startup	PRISE Framework
Core Objective	User experience innovation	Business model validation	Technology value realization
Risk Focus	Requirement deviation	Market misalignment	Technical feasibility risk
Validation Object	Prototype usability	Business hypothesis	Technical value realizability
AI Adaptation Weakness	Neglects technical feasibility	Ignores data dependency	No inherent weaknesses

metrics → business metrics, ensuring AI capability improvements truly translate into business value. For example, creating clear associations between dialogue system accuracy improvements and user satisfaction and conversion rates forms a complete value chain from technology to behavior to business.

- **Engineering Genes**: Simultaneously designs production-ready capabilities such as monitoring and rollback in the end-to-end productization stage, complementing lean startup's deficiencies in technical delivery aspects, ensuring AI products can provide stable service as long-term operating systems.

Through these three innovative characteristics, the PRISE framework provides more comprehensive, systematic guidance for AI products from concept to implementation. It retains the user-centered perspective of design thinking and the rapid validation concept of lean startup while adding technology support mechanisms specific to AI products, forming a methodological system suitable for the entire lifecycle of AI product incubation.

5) PRISE FRAMEWORK SUCCESS METRICS AND QUANTITATIVE INDICATORS

To enhance the operational effectiveness and measurable outcomes of the PRISE framework, we establish specific quantitative indicators for each dimension. These metrics provide objective assessment criteria for practitioners to evaluate progress and identify areas for improvement during AI product incubation. Table 3 presents the comprehensive measurement framework across all five PRISE dimensions.

These quantitative indicators serve multiple purposes: (1) providing objective benchmarks for team performance assessment; (2) enabling comparative analysis across different AI product initiatives; (3) facilitating continuous improvement through measurable feedback; and (4) supporting resource allocation decisions based on empirical evidence. The threshold values are derived from industry best practices and validated through our cross-case analysis, representing achievable yet challenging targets for AI product development teams.

III. RESEARCH METHODOLOGY

This research employs an explanatory multiple-case study method [28] to validate the PRISE framework through seven cross-industry AI product incubation cases. Cases were selected through theoretical sampling from 27 candidate projects, meeting criteria of: complete concept-to-launch experience, traceable decision points, comprehensive documentation, and available key stakeholders. The development of the PRISE framework followed a systematic mixed-method approach combining theoretical synthesis with empirical induction, ensuring both theoretical rigor and practical validity.

A. FRAMEWORK DEVELOPMENT PROCESS

The PRISE framework emerged through a structured four-phase process integrating literature insights with case-derived patterns:

Phase 1 - Theoretical Foundation: Systematic literature review of academic research and industry reports on product development methodologies identified three critical gaps when applied to AI products: insufficient early technical feasibility assessment, weak value verification mechanisms, and inadequate engineering transformation support.

Phase 2 - Exploratory Analysis: In-depth analysis of pilot cases using open coding revealed five recurring patterns: upfront problem-technology integration, rapid validation adaptation, value transformation pathways, vertical focus effectiveness, and early engineering integration requirements. These patterns emerged inductively from case observation rather than deductive theoretical reasoning.

Phase 3 - Framework Conceptualization: The empirical patterns were systematically integrated with established innovation theories to form the five-dimensional PRISE structure, ensuring both empirical grounding and theoretical foundations.

Phase 4 - Validation and Refinement: Additional cross-industry cases validated the framework's explanatory power and practical applicability, with expert review from AI product professionals informing final refinements.

B. DATA COLLECTION AND ANALYSIS

Data collection employed triangulation across multiple sources: 21 semi-structured interviews with project stakeholders, 7 focus group discussions, participatory observation, and comprehensive project documentation. Analysis followed a dual approach: open coding for pattern discovery in exploratory cases and focused coding based on PRISE dimensions for validation cases. Two researchers

TABLE 3. PRISE framework quantitative success metrics.

Dimension	Metric Level	Quantitative Indicator	Target Threshold	Description
Problem Definition	Core Metric	Requirements First-Pass Rate	$\geq 70\%$	Success rate of AI team's initial requirements confirmation with business stakeholders
	Supporting Metric	Interview Role Coverage	≥ 2 roles	Minimum stakeholder diversity in user research
	Supporting Metric	Requirement Iteration Rounds	≤ 3 rounds	Maximum communication cycles for requirement finalization
	Supporting Metric	Definition Phase Duration	≤ 30 days	Time constraint for problem definition completion
Rapid Validation	Core Metric	Validation Cycle Duration	≤ 14 days	Maximum time for hypothesis validation completion
Impact Evaluation	Core Metric	Successful Hypothesis Ratio	$\geq 60\%$	Minimum success rate for validated hypotheses
	Supporting Metric	Estimated ROI	$> 150\%$	Expected return on investment threshold
Single-focus Development	Core Metric	MVP Launch Cycle	≤ 90 days	Time from concept to minimum viable product launch
	Supporting Metric	A/B Test Statistical Significance	Positive business KPI improvement	Required outcome for controlled experiments
End-to-End Productization	Core Metric	Service Uptime Rate	$\geq 99.5\%$	Long-term system availability requirement
	Core Metric	Critical Incidents (30 days)	0 incidents	Maximum acceptable severe failures post-launch
	Supporting Metric	Monitoring Coverage Rate	$\geq 90\%$	Percentage of system components under monitoring

independently coded all data, achieving 83% inter-rater reliability.

C. RESEARCH QUALITY ASSURANCE

Quality assurance addressed four validity dimensions: construct validity (case database establishment, data triangulation, key informant review), internal validity (pattern matching, explanatory frameworks, alternative analysis), external validity (multiple-case design, theoretical sampling, theory comparison), and reliability (standardized protocols, systematic documentation). These measures ensured the scientific rigor and credibility of findings.

IV. CASE VALIDATION: PRACTICAL APPLICATION OF THE PRISE FRAMEWORK

This section validates the practical application value and effects of the PRISE framework in the AI product incubation process through seven cross-industry cases. The analysis unfolds according to the framework's five dimensions, highlighting the core value and application patterns of each dimension.

A. CASE OVERVIEW

The seven cases analyzed in this study cover multiple industries including education, enterprise services, healthcare, and retail, as shown in Table 4.

B. PROBLEM DEFINITION DIMENSION

The Problem Definition dimension focuses on user role identification and needs understanding, forming the foundational stage of AI product incubation.

1) CASE 1: AIRLEARNING USER ROLE MISJUDGMENT

AirLearning was an AI music education product for K-12 students with excellent technical performance but low paid conversion rates. Research revealed the core problem was incorrectly assuming children were the decision makers, while parents actually controlled purchasing decisions. The product's correction feedback, though technically accurate, created frustration for children, reducing parents' willingness to pay. After re-identifying the "family music learning decision chain," the team shifted to parent-child co-experience design, optimized feedback mechanisms, and emphasized progress visualization, significantly improving both conversion rates and retention.

2) CASE 2: ACCELERATE EXCELLENCE PROJECT GAP

Microsoft's sales platform project team spent months collecting extensive research data through interviews and white paper analyses but couldn't translate it into a clear product direction. By applying the "technical anchoring mechanism" and "follow-along research," they discovered sales personnel

TABLE 4. Case overview and PRISE dimension mapping.

Case	Case Title	Industry Background	Core Challenge	Primary PRISE Dimension
Case 1	AirLearning User Role Misjudgment	Educational Music App	User identification error	Problem Definition
Case 2	Accelerate Excellence Project Gap	Enterprise Sales AI	Research-practice disconnect	Problem Definition
Case 3	Dialog Engine Value Expression	Dialogue Platform	Unclear value communication	Impact Evaluation
Case 4	Rapid Validation Culture Building	Internal Innovation Mechanism	Missing validation process	Rapid Validation
Case 5	E-commerce Platform Portal A/B Testing	E-commerce Platform	User experience optimization	Rapid Validation
Case 6	AI Shopping Assistant Value Insufficiency	Retail Shopping Assistant	Weak user value perception	Impact Evaluation
Case 7	Healthcare AI Vertical Focus	Healthcare Solution	Resource dispersion	Single-focus Development

spent substantial time daily integrating information across multiple systems. The team focused on developing an intelligent assistant for specific scenarios, saving sales personnel significant working hours weekly. The case demonstrates that deeply understanding work processes yields more effective product insights than broad theoretical research.

3) KEY FINDINGS IN THE PROBLEM DEFINITION DIMENSION

- Precise identification of user roles and decision chains is prerequisite for AI product success, particularly distinguishing between users and decision makers
- Theoretical research must combine with daily workflow observation to obtain implementable product insights
- Technical anchoring mechanisms can assess solution technical feasibility early in problem definition, avoiding later resource waste

C. RAPID VALIDATION DIMENSION

The Rapid Validation dimension emphasizes validating key hypotheses at minimum cost, ensuring correct product direction and avoiding unnecessary development investments.

1) CASE 4: RAPID VALIDATION CULTURE BUILDING

Microsoft's AI vertical team faced lengthy decision cycles for new ideas. After introducing the concept that "all ideas are bad by default," they established a systematic validation mechanism: (1) breaking down core hypotheses, (2) designing low-cost validation experiments, (3) setting clear success criteria, and (4) making decisions quickly. The new mechanism shortened decision cycles significantly, completed numerous idea validations efficiently, and avoided substantial resource waste. One AI content generation tool validated core hypotheses in just days using the Wizard of Oz method, saving months of development time.

2) CASE 5: E-COMMERCE PLATFORM PORTAL A/B TESTING

After integration with a cross-border e-commerce platform, Microsoft's advertising team found merchant activation rates were low. The team designed a simple A/B test, moving the advertising entry point to the merchant backend homepage, significantly increasing the test group activation rate. This straightforward experiment directly raised the overall activation rate and prompted the team to establish continuous optimization mechanisms. The case proves that lightweight experiments can quickly solve complex problems, avoiding large-scale development risks.

3) KEY FINDINGS IN THE RAPID VALIDATION DIMENSION

- Systematic hypothesis validation mechanisms can significantly shorten decision cycles and improve resource utilization efficiency
- Low-cost validation methods (such as Wizard of Oz, A/B testing) are particularly suitable for AI product prototype validation
- "Data-driven decision making" culture is more effective than "opinion-driven decision making," reducing subjective arguments

D. IMPACT EVALUATION DIMENSION

The Impact Evaluation dimension focuses on transforming AI technical capabilities into user-perceived value and business value, resolving the "technology-to-value" conversion challenge in AI products.

1) CASE 3: DIALOG ENGINE VALUE EXPRESSION

Dialog Engine was a Chinese AI dialogue platform with excellent technical performance but low sales conversion. Research found that customers recognized technical capabilities but were unclear about business value. After applying the "value assessment canvas" tool, the team reconstructed the value expression framework, establishing clear associations between technical metrics and business value (e.g., NLU accuracy → improvement in first-time resolution rate → increase in customer satisfaction → reduction in customer service costs), and developed an ROI calculator to help customers assess returns. The sales conversion rate increased significantly within months, approaching target levels.

2) CASE 6: AI SHOPPING ASSISTANT VALUE INSUFFICIENCY

The retail industry AI shopping assistant was technically ‘‘good enough’’ (answering questions, supporting multi-turn dialogues), but user engagement was low in testing. Comparative task analysis showed that completing shopping tasks took longer with the AI assistant than with existing search functionality. The team redefined the product's differentiated value, focusing on unique AI capabilities (visual recognition, outfit recommendations). After revision, conversation turns increased substantially, and most users expressed willingness to use it again.

3) KEY FINDINGS IN THE IMPACT EVALUATION DIMENSION

- AI products must establish clear transformation pathways from technical metrics → behavioral metrics → business metrics
- For new AI products, ‘‘good enough’’ is not enough; they must provide significantly better differentiated value than existing solutions
- Value communication should focus on outcomes that users/customers care about, rather than technical capabilities themselves

E. SINGLE-FOCUS DEVELOPMENT DIMENSION

The Single-focus Development dimension emphasizes deeply cultivating a single vertical scenario with limited resources, quickly establishing successful cases and a foundation of trust.

1) CASE 7: HEALTHCARE AI VERTICAL FOCUS

a: BACKGROUND AND CHALLENGES

Microsoft's AI Vertical team initially attempted to advance AI solutions across multiple industries (retail, finance, healthcare, manufacturing). Despite strong technical capabilities, dispersed resources led to slow progress—they contacted numerous clients across various industries but failed to implement a single successful commercialized solution. The team leader candidly stated: ‘‘The breadth we pursued prevented us from achieving sufficient depth in any area, a typical ‘resource dispersion effect’ problem.’’

b: PROBLEM AND PROCESS

After strategic assessment, the team applied the ‘‘vertical scenario evaluation tool’’ from the PRISE framework, systematically evaluating opportunities across industries based on three dimensions (value potential, technical feasibility, and data availability). Results showed healthcare leading significantly in the comprehensive score.

The team decided to focus on the healthcare vertical domain for reasons including: (1) high degree of healthcare data structuring; (2) clear industry pain points with high value; (3) team technical advantages in medical NLP; (4) although healthcare market access is difficult, barriers are high once entered. The team's AI architect explained: ‘‘To establish a product firmly, you need early adopters who can become your true supporters or advocates. Therefore, we chose a vertical domain where we could build deep impact, rather than superficially dabbling in multiple areas.’’

The team invested all resources in the healthcare field, focusing on three core scenarios: medical entity recognition, clinical text structuring, and intelligent assisted diagnosis. The product manager emphasized: ‘‘We implemented a ‘subtraction strategy,’ considering not what we could do, but what we shouldn't do, focusing limited resources on areas that could truly create breakthroughs.’’

c: IMPLEMENTATION AND RESULTS

After implementing the focused strategy, the team developed several high-impact healthcare AI use cases and established deep partnerships with top-tier hospitals. The first successful case—a medical imaging assisted diagnosis system—increased doctor diagnostic efficiency and improved diagnostic accuracy in the pilot hospital.

These early successful samples quickly established the team's reputation and trust in the healthcare AI field, creating a flywheel effect, attracting more partners and increased resource support. A hospital radiology department director stated: ‘‘Microsoft's AI team's depth of understanding of the healthcare field is impressive. They don't simply apply generic AI technology but truly understand the professional needs of medical image diagnosis.’’

Later, the team had become a leader in the healthcare AI field, with products covering numerous hospitals. The team leader reflected: ‘‘Focusing on a single vertical domain was our most critical strategic decision, transforming us from ‘casting nets everywhere yet catching nothing’ to ‘deep cultivation and building influence.’’’

2) KEY FINDINGS IN THE SINGLE-FOCUS DEVELOPMENT DIMENSION

- ‘‘Wide net’’ strategies are ineffective in AI product development; concentrating resources in one vertical domain more easily achieves breakthroughs
- Scenario selection should comprehensively consider value potential, technical feasibility, and data availability
- Quickly establishing successful cases in small but focused scenarios is more effective than attempting to build comprehensive solutions

F. END-TO-END PRODUCTIZATION DIMENSION

The End-to-End Productization dimension focuses on how AI products successfully transition from concept prototypes to production environments, resolving the challenge of AI model engineering implementation.

1) PRODUCTIZATION EXPERIENCES FROM CASES 5 AND 6

Both the e-commerce platform portal project and AI shopping assistant case demonstrate the importance of the End-to-End Productization dimension. The e-commerce platform project rapidly integrated A/B testing results into the product development process, establishing continuous optimization mechanisms that steadily increased activation rates. The AI shopping assistant project redesigned the end-to-end delivery process from a user experience perspective, ensuring new features (such as visual recognition) could seamlessly integrate into the user shopping journey. Both cases prove that AI products need to consider end-to-end user experience and production environment adaptability from the conceptual stage, rather than focusing solely on technical function implementation.

2) KEY FINDINGS IN THE END-TO-END PRODUCTIZATION DIMENSION

- AI product engineering must consider reliability, maintainability, and scalability in production environments
- End-to-end user experience design is crucial for AI product adoption
- Early establishment of feedback loops and continuous optimization mechanisms can significantly enhance AI product long-term value

G. CROSS-CASE ANALYSIS: DERIVATION OF AI PRODUCT DESIGN PRINCIPLES

Through systematic analysis of seven cases, we derived three transferable design principles that constitute the theoretical foundation for AI product incubation. These principles, grounded in organizational learning theory [2], [18] and innovation heuristics [24], transform case-specific observations into actionable guidance applicable across diverse contexts.

1) DESIGN PRINCIPLE 1: STAKEHOLDER ROLE DIFFERENTIATION

Principle Statement: AI product teams must systematically distinguish between end users, influencers, and decision makers, designing distinct value propositions for each stakeholder category to ensure alignment between product capabilities and decision-making processes.

Theoretical Foundation: This principle builds on stakeholder theory and multi-level decision-making frameworks, recognizing that AI products often operate in complex organizational contexts where usage patterns and purchasing decisions are distributed across different roles.

Operational Guidelines:

- Conduct stakeholder mapping exercises early in the problem definition phase
- Develop role-specific value narratives that address distinct stakeholder concerns
- Design user experiences that satisfy end-user needs while facilitating decision-maker evaluation
- Establish feedback loops that capture insights from all stakeholder categories

Empirical Evidence: Cases 1 (AirLearning) and 3 (Dialog Engine) demonstrate that failure to apply this principle results in product designs that satisfy technical requirements but fail in market adoption due to misaligned stakeholder incentives.

2) DESIGN PRINCIPLE 2: HYPOTHESIS-DRIVEN RESOURCE OPTIMIZATION

Principle Statement: AI product development should prioritize low-cost hypothesis validation over comprehensive feature development, systematically testing core assumptions before committing significant resources to implementation.

Theoretical Foundation: Rooted in lean thinking and bounded rationality principles, this approach recognizes the high uncertainty inherent in AI product development and advocates for systematic risk reduction through iterative learning.

Operational Guidelines:

- Decompose product concepts into testable hypotheses ranked by risk and importance
- Design minimum viable experiments that provide maximum learning per resource unit invested
- Establish clear success criteria and decision thresholds before conducting validation experiments
- Create organizational processes that reward rapid learning over comprehensive planning

Empirical Evidence: Cases 4 (Rapid Validation Culture) and 5 (E-commerce Platform A/B Testing) illustrate how systematic application of this principle accelerates decision cycles and improves resource allocation efficiency.

3) DESIGN PRINCIPLE 3: VALUE TRANSLATION TRANSPARENCY

Principle Statement: AI products must establish explicit pathways connecting technical performance metrics to user-perceivable benefits and business outcomes, ensuring that technological advances translate into stakeholder value.

Theoretical Foundation: This principle addresses the "black box" problem in AI systems by creating interpretable value chains that bridge technical capabilities and business impact, drawing on theories of technology acceptance and value co-creation.

Operational Guidelines:

- Develop multi-level metrics frameworks linking technical → behavioral → business indicators
- Create value demonstration tools that translate technical improvements into stakeholder-relevant outcomes
- Establish continuous measurement systems that track value realization across the entire chain
- Design communication strategies that articulate value propositions in stakeholder-appropriate language

Empirical Evidence: Cases 3 (Dialog Engine) and 6 (AI Shopping Assistant) show that systematic application of this principle transforms technically excellent products into commercially viable solutions.

4) INTEGRATION AND TRANSFER MECHANISMS

These design principles are interrelated and mutually reinforcing, forming a coherent approach to AI product development that addresses the unique challenges identified in our theoretical analysis. The principles are designed to be context-adaptable, providing general guidance that can be customized for specific industries, organizational structures, and technological constraints.

The transferability of these principles is supported by their grounding in established organizational learning and innovation management theories, while their specificity to AI products is validated through empirical evidence from diverse industry contexts. This combination ensures both theoretical rigor and practical applicability across different AI product development scenarios.

V. PRISE FRAMEWORK APPLICATION AND VALIDATION

This section articulates the PRISE framework's application methods, validation mechanisms, and applicable boundaries, providing systematic operational guidance for practitioners.

A. PRISE FRAMEWORK APPLICATION GUIDE

The PRISE framework provides practical guidance in both team management and product design dimensions:

1) TEAM MANAGEMENT ACTIVITIES

- **Problem Definition:** Form cross-functional teams, establish technical feasibility assessment
- **Rapid Validation:** Build "fail fast" culture, establish hypothesis repositories
- **Impact Evaluation:** Establish three-layer metric systems, hold value reviews
- **Single-focus:** Implement "subtraction thinking," set clear phased goals
- **Productization:** Form complete teams early, establish monitoring mechanisms

2) PRODUCT DESIGN TOOLS

- **Problem Definition:** User role maps, decision chain analysis, feasibility matrices
- **Rapid Validation:** Validation cards, Wizard of Oz method, MVP experiment design
- **Impact Evaluation:** Value assessment canvas, A/B testing framework, ROI calculator
- **Single-focus:** Scenario evaluation tool, user journey design, MVP definition
- **Productization:** Readiness checklist, performance monitoring, user guidance

Practical tools developed in this research include feasibility matrices, value assessment canvases, hypothesis validation cards, vertical scenario evaluation tools, and productization readiness checklists, providing structured methodological support for each stage.

B. PRISE FRAMEWORK VALIDATION MECHANISMS

Framework validation employs a three-layer mechanism:

Case Application Validation: Examines the performance of each framework dimension in actual projects through seven cross-industry cases, verifying whether the PRISE framework can effectively solve key challenges in AI product incubation. Analysis focuses include team decision-making process changes, solution quality improvements, and project progress rhythm enhancements.

Project Progress Assessment: Evaluates the framework's impact on various stages of the project lifecycle, including proof-of-concept acceleration, iteration cycle changes, decision efficiency improvements, focusing on how the framework changes team working modes and delivery capabilities.

External Expert Review: Invites AI product experts from academia and industry to systematically evaluate the framework, providing assessments on theoretical completeness, operational feasibility, universal applicability, and value creation potential, ensuring the framework's scientific nature and practicality.

C. FRAMEWORK APPLICABLE CONDITIONS AND BOUNDARIES

The PRISE framework is not universally applicable. Table 5 summarizes its application scenarios and limitations.

Practitioners should flexibly apply the PRISE framework according to specific project characteristics, making scenario-based adjustments when necessary to maximize its guidance value.

D. PRISE LITE: FRAMEWORK ADAPTATION FOR ORGANIZATIONAL SCALE AND RESOURCE CONSTRAINTS

Recognizing that organizations vary significantly in scale, resource availability, and governance maturity, the complete PRISE framework may not be suitable for all contexts. To address this limitation, we introduce PRISE Lite, a streamlined adaptation designed for resource-constrained environments and early-stage initiatives.

1) ORGANIZATIONAL AND RESOURCE CLASSIFICATION FRAMEWORK

We classify organizations across two dimensions: organizational scale and resource capability level. Organizational scale considers industry-specific thresholds—for instance, in China's software and information technology services sector: small enterprises (<100 employees), medium enterprises (100-300 employees), and large enterprises (> 300 employees). Resource capability encompasses technical resources, data governance maturity, and expected AI project investment, as detailed in Table 6.

TABLE 5. PRISE framework application boundaries.

Category	Characteristics
Most Suitable Scenarios	• Enterprise-level AI solutions, especially B2B and B2B2C applications • Products dependent on large volumes of structured or unstructured data • Highly uncertain innovation projects from concept to implementation
Potentially Unsuitable Scenarios	• Pure research-oriented AI technology exploration projects • Application areas with mature product models available for reference • Very small-scale personal development projects
Application Limitations	• Requires interdisciplinary team support (product, technology, business) • Some tools require training and experience accumulation • More suitable for medium to long-term projects (3+ months)

TABLE 6. Resource capability classification framework.

Resource Level	Technical Resources	Data Governance	Budget Range
Basic	No software/AI development capabilities, limited deployment capacity, cannot independently establish databases or deploy software systems	Low digitization, scattered data with minimal governance, no metadata or labeling systems	$\leq$ \$75K
Standard	Basic software deployment and operations capability, low/no-code development capacity	Most data digitized with some governance, but departmental silos exist without unified strategy	\$75K - \$450K
Advanced	Software development capabilities, some AI development capacity	Complete data digitization with unified governance strategy and productized data systems	$\geq$ \$450K

TABLE 7. PRISE framework selection matrix.

Resource Level	Small Org	Medium Org	Large Org
Basic	PRISE Lite	PRISE Lite	PRISE Lite
Standard	PRISE	PRISE	PRISE Lite
Advanced	PRISE	PRISE	PRISE

The framework selection matrix combines organizational scale and resource levels to determine the appropriate approach, as shown in Table 7.

Notably, large organizations with Standard resource levels are recommended to use PRISE Lite due to their complex organizational structures and higher cross-departmental collaboration challenges, which can hinder full framework implementation despite adequate resources.

2) PRISE LITE FRAMEWORK SPECIFICATION

PRISE Lite represents a focused subset of the complete framework, designed to answer two critical questions with minimal investment: "Can we build it?" and "Should we build it?" The framework comparison is presented in Table 8.

PRISE Lite enables organizations to establish a foundation for AI product development while building internal capabilities and demonstrating value, creating a pathway for eventual graduation to the complete PRISE framework as resources and organizational maturity increase.

E. ETHICS AND COMPLIANCE CONSIDERATIONS

Future implementations of the PRISE framework should systematically integrate ethical considerations and regulatory compliance throughout the development lifecycle. Key areas for enhanced ethical integration include:

Algorithmic Fairness: Each PRISE dimension should incorporate bias assessment mechanisms, ensuring equitable outcomes across different user demographics and use cases.

Privacy Protection: Data collection and processing activities must align with evolving privacy regulations, requiring privacy-by-design principles embedded within the technical anchoring and validation processes.

Transparency and Explainability: AI products developed using PRISE should maintain interpretability standards appropriate to their deployment context, particularly in high-stakes applications.

Human Oversight: The framework should emphasize meaningful human control mechanisms, ensuring that automation enhances rather than replaces human judgment in critical decisions.

These ethical considerations represent evolving requirements that will likely require continuous framework adaptation as regulatory landscapes and societal expectations develop.

VI. DISCUSSION AND OUTLOOK

This section discusses the research findings, theoretical contributions, practical implications, limitations, and future research directions of the PRISE framework.

A. RESEARCH FINDINGS AND THEORETICAL CONTRIBUTIONS

This research demonstrates the effectiveness of the PRISE framework across multiple AI product incubation cases,

TABLE 8. PRISE vs. PRISE lite comparison.

Dimension	PRISE Lite	Key Differences from Full PRISE
Problem Definition	Business pain point mapping, decision chain analysis, data landscape assessment, technical anchoring (data/compute availability scoring)	Similar scope but narrower focus, reduced breadth of investigation
Rapid Validation	POC design with 3-5 key metrics, open-source tool implementation, synthetic or sampled data usage	Full PRISE requires real environment data, multiple prototype iterations, and user segmentation experiments
Impact Evaluation	Verify POC metrics, estimate preliminary ROI through technical-to-business transformation, no long-term behavioral modeling	Full PRISE requires three-layer metric chains and small-scale production validation
Single-focus Development	Optional. If POC succeeds, proceed to small-scale single-scenario delivery; otherwise, end Lite cycle	Full PRISE requires mandatory implementation with broader scope
End-to-End Productization	Not covered (provides next-step engineering checklist only)	Full PRISE includes complete architecture, monitoring, rollback design, and production deployment

revealing three significant theoretical contributions that extend and complement established innovation management theory:

1) BRIDGING THE AI-SPECIFIC "VALLEY OF DEATH" THROUGH TECHNICAL ANCHORING

Our first contribution addresses the classic innovation challenge known as the "valley of death"—the gap between research and practical application [16]—in the context of AI innovation. Traditional innovation theories emphasize either technology-push or market-pull approaches [26], but AI innovation requires simultaneous consideration of both dimensions due to unique data dependencies and algorithmic constraints.

The PRISE framework's "technical anchoring mechanism" represents a novel solution to this challenge, establishing clear technical feasibility boundaries during early problem exploration. This approach prevents the common AI innovation failure pattern where technically sophisticated solutions fail due to insufficient data quality or unrealistic performance expectations. Case 7 (Healthcare AI) demonstrates how this mechanism substantially reduces the traditional valley of death width, significantly reducing prototype iteration cycles and accelerating time-to-market. This finding extends classical innovation failure theory by identifying AI-specific failure modes and providing systematic mitigation strategies.

2) ENHANCING TECHNOLOGY ADOPTION THROUGH MULTI-LAYER VALUE VERIFICATION

Our second contribution builds upon Technology Acceptance Model (TAM) and Diffusion of Innovations theory [21], [27] by addressing the unique trust and explainability challenges in AI adoption. Traditional adoption models focus on perceived usefulness and ease of use, but AI systems introduce additional complexity through algorithmic opacity and uncertain performance boundaries.

The PRISE framework's "value verification closed loop" systematically addresses these challenges through a three-layer validation model (technical feasibility, business feasibility, market acceptance) that explicitly connects technical performance to user-perceived value. This approach operationalizes the sequential logic of Rogers' innovation-decision process within the AI product context [21], and aligns with the Minimum Marketable Features (MMF) concept by emphasizing early, iterative value delivery. Case 2 (Intelligent Content Generation) illustrates how this multi-layer approach helps teams navigate Christensen's innovation paradox [4] by identifying sustainable value propositions that balance technical capability with market needs.

Furthermore, our findings contribute to the emerging discourse on AI trust and explainability [9], [20] by demonstrating how systematic value translation mechanisms can enhance user acceptance of AI systems, even when the underlying algorithms remain opaque.

3) RESOLVING THE TECHNOLOGY COMMERCIALIZATION CHALLENGE THROUGH EARLY ENGINEERING INTEGRATION

Our third contribution addresses the persistent challenge of technology commercialization—transitioning from proof-of-concept to scalable production systems. This challenge is particularly acute in AI development, where industry research [6] shows that over 60% of AI proof-of-concept projects fail to reach production environments.

The concept of "early engineering gene implantation" represents a fundamental departure from traditional stage-gate innovation models by embedding production considerations into the design phase, rather than as subsequent activities. This approach draws inspiration from concurrent engineering principles but adapts them specifically for AI systems' unique requirements including data quality management, model explainability, and continuous monitoring capabilities.

Our empirical evidence demonstrates that this approach significantly improves prototype-to-production transformation success rates, effectively eliminating a major bottleneck in AI innovation. This finding contributes to innovation management theory by providing empirical validation of

TABLE 9. Value assessment canvas template.

Item	Component	Description	Example
1.1	**User Pain Point**	Specific business pain point for a particular role in a defined scenario	Customer service agents spend excessive time on routine inquiries
1.2	**Target AI Capability**	Technical metrics and expected outcomes	Intent classification precision improvement by 2%
1.3	**Direct Business Metric**	Business KPI and current baseline	Average handling time, current baseline: 180 seconds
1.4	**Technical-to-Business Transformation Hypothesis**	Formula linking technical metrics to business KPIs	Precision +1% $\Rightarrow$ handling time -5 seconds
1.5	**Expected Benefits**	Anticipated revenue increase, efficiency gains, or risk reduction	15% reduction in operational costs
1.6	**Validation Method**	A/B testing, Wizard of Oz, simulation, etc.	A/B test with 30% traffic allocation
1.7	**Confidence Score**	Confidence level (high/medium/low) with rationale	Medium - based on similar industry implementations

early-stage engineering integration benefits and offering a systematic framework for implementation.

4) THEORETICAL INTEGRATION AND EXTENSIONS

These three contributions collectively extend classical innovation theory in several important ways. First, they provide empirical evidence for the necessity of AI-specific innovation frameworks that cannot be adequately addressed by adapting traditional methodologies. Second, they demonstrate how systematic integration of technical, business, and engineering considerations can significantly improve innovation success rates in complex technological domains. Third, they offer practical guidance for navigating Moore's "chasm" [15] in AI product development by providing concrete mechanisms for building market readiness and user adoption.

The PRISE framework thus represents both a practical tool for AI innovation practitioners and a theoretical extension of innovation management theory into the AI domain, providing new insights into how traditional innovation challenges manifest and can be addressed in technology-intensive, data-dependent innovation contexts.

B. PRACTICAL IMPLICATIONS

Based on research findings, this study provides the following key implications for AI product development practice:

Dual Anchoring of Requirements and Technology: AI product teams should simultaneously value user needs and technical capability boundaries, clarifying "what technology won't do" at the product vision stage, avoiding falling into unlimited expansion and technical perfectionism.

Multi-level Value Hypothesis Verification: Verification should cover technical feasibility, business model feasibility, and market acceptance, adopting the concept of "minimum viable dataset" for rapid iterative validation.

Front-loaded Engineering Awareness: Product teams should consider data quality, algorithm maintainability, and system scalability from the concept stage, ensuring smooth transition to scaled deployment.

Cross-functional Collaboration Mechanism: Establish regular collaboration mechanisms among product, technology, and business parties, evaluating product iteration directions with a "value-first, cost-second" principle.

C. RESEARCH LIMITATIONS AND CASE SCOPE ANALYSIS

This research acknowledges several important limitations that should be considered when interpreting the findings and applying the PRISE framework in different contexts:

Geographic and Cultural Scope Limitations: The seven case studies are primarily derived from Microsoft's operations in China and the United States, with limited representation from other geographic regions including Europe, Latin America, Africa, and Southeast Asia. This geographic concentration may limit the framework's cross-cultural applicability, as different regions have varying approaches to technology adoption, regulatory environments, organizational decision-making processes, and innovation cultures. For instance, European markets place stronger emphasis on privacy and regulatory compliance, which may require framework adaptations beyond those described in our ethics and compliance section.

Industry Sector Limitations: While our cases span multiple industries (education, healthcare, retail, enterprise services), they are primarily concentrated in the technology and digital services sectors. The framework's applicability to traditional industries with limited digital transformation experience, such as manufacturing, agriculture, or construction, requires additional validation. Furthermore, our cases focus predominantly on B2B and B2B2C applications, with limited insight into B2C consumer applications or government/public sector AI implementations.

TABLE 10. Validation card template.

Field	Content	Example
ID	Unique hypothesis identifier	S-V-003
Description	Clear statement of the hypothesis being tested	Placing ad entry on merchant dashboard homepage increases activation rate
Validation Metric	Primary metric for success measurement	Activation Rate
Success Criteria	Statistical threshold and significance level	Activation Rate +15%, $p < 0.05$
Experiment Design	Detailed experimental approach	A/B test with 50% traffic over 7 days
Resource Investment	Required resources for validation	1 developer-day, 50% traffic allocation
Validation Result	Outcome checkboxes	☐ Success ☐ Failure ☐ Requires iteration
Next Actions	Follow-up steps for each possible outcome	If successful: proceed to MVP; If failed: conduct retrospective

TABLE 11. Vertical scenario evaluation framework.

Dimension	Score Range	Weight	Evaluation Criteria
Value Potential	1-5	0.4	**1-2:** Small/unclear market, weak pain points, low willingness to pay **3-4:** Medium market, noticeable pain points, moderate willingness to pay **5:** Large and growing market, strong pain points, high willingness to pay
Technical Feasibility	1-5	0.3	**1-2:** Critical data missing/algorithms immature, requires extensive R&D **3-4:** Data generally available, mainstream algorithms applicable, moderate integration work **5:** Abundant and accessible data, mature algorithms, rapid deployment possible
Data Availability	1-5	0.2	**1-2:** Mostly unstructured or unlabeled data, low quality **3-4:** Partially structured, incomplete labels, requires additional cleaning/annotation **5:** Highly structured data, complete labels, established governance
Entry Barriers	1-5	0.1	**1-2:** High compliance requirements, significant industry barriers, long cycles **3-4:** Manageable compliance, moderate barriers, requires qualifications/partnerships **5:** Low compliance requirements, minimal barriers, rapid market entry possible

Organizational Context Limitations: All case studies derive from Microsoft's AI initiatives, representing a large technology corporation with substantial resources, technical expertise, and established AI capabilities. This organizational context may not reflect the challenges faced by startups, small-to-medium enterprises, non-profit organizations, or companies in regions with less developed AI ecosystems. While our PRISE Lite framework addresses some resource constraints, broader organizational diversity is needed for comprehensive validation.

Temporal and Technology Evolution Limitations: Our research covers AI product development projects primarily conducted between 2020-2024, focusing on machine learning applications including natural language processing, computer vision, and predictive analytics. The rapid evolution of AI technology, particularly the emergence of large language models and generative AI capabilities, may require framework adaptations not fully captured in our current analysis. Additionally, the relatively short observation period limits our understanding of long-term product evolution and market sustainability.

Methodological Limitations: The case study methodology, while providing rich contextual insights, inherently limits generalizability compared to large-scale quantitative studies. Our reliance on retrospective analysis and participant recollection may introduce memory bias, although we mitigated this through multiple data sources and triangulation. The framework's effectiveness metrics are based on project-specific indicators rather than standardized industry benchmarks, which may limit cross-case comparability.

Stakeholder Perspective Limitations: Our research primarily captures perspectives from internal development teams, product managers, and organizational leadership. We have limited representation of end-user perspectives, external stakeholders, regulatory bodies, or broader societal impacts. This limitation is particularly relevant for understanding the social implications of AI products and the effectiveness of our proposed ethical frameworks.

Implications for Framework Application: These limitations suggest that practitioners should adapt the PRISE framework to their specific contexts rather than applying it prescriptively. Organizations operating in different geographic regions, industry sectors, or organizational

TABLE 12. Productization readiness checklist.

Category	Assessment Item	Ready Criteria
Technical Architecture	Model performance meets production thresholds	Accuracy > 85%, Latency < 200ms
	Scalability testing completed	Handles 10x expected load
	Error handling mechanisms implemented	Graceful degradation, fallback options
	Security assessment passed	Data encryption, access control validated
	API documentation complete	Full endpoint documentation with examples
	Integration testing successful	End-to-end workflow validation
Data & Monitoring	Data pipeline established	Automated data ingestion and processing
	Model monitoring deployed	Performance drift detection active
	Logging system operational	Comprehensive request/response logging
	Alerting mechanisms configured	Real-time performance and error alerts
	Data governance compliance verified	Privacy, retention policies implemented
Operational Readiness	Deployment automation configured	CI/CD pipeline for model updates
	Rollback procedures tested	Version rollback within 15 minutes
	Disaster recovery plan validated	Backup systems and recovery procedures
	Performance benchmarks established	Baseline metrics and SLA definitions
	Support documentation created	Troubleshooting guides and runbooks
	Team training completed	Operations team familiar with system
Business Integration	User acceptance testing passed	Representative user validation completed
	Business metrics defined	ROI tracking and success criteria
	Compliance requirements met	Regulatory and legal requirements verified
	Change management plan ready	User training and adoption strategy
	Support processes established	Customer support and escalation paths
	Go-to-market strategy finalized	Launch timeline and communication plan

contexts should conduct pilot implementations and gather local feedback before full-scale adoption. Additionally, the framework should be viewed as a starting point for developing context-specific AI product development approaches rather than a universal solution.

D. EMERGING AI TECHNOLOGY ADAPTATIONS

The rapid evolution of AI technology, particularly large language models and generative AI, necessitates framework adaptations to address new development paradigms:

Foundation Model Integration: The availability of pre-trained foundation models transforms the technical anchoring mechanism, requiring assessment of model selection versus custom development, prompt engineering complexity, and computational resource requirements.

Generative AI Considerations: Generative systems introduce unique validation challenges including output quality consistency, bias in generated content, and alignment with intended use cases. The framework's rapid validation dimension must accommodate testing methodologies specific to generative capabilities.

Autonomous System Requirements: AI agents capable of autonomous decision-making require enhanced safety considerations, human oversight mechanisms, and trust calibration strategies that extend beyond traditional AI application requirements.

Technology Stack Evolution: Modern AI products increasingly integrate multiple AI technologies, creating complexity in system design, version management, and performance optimization that the framework must address through flexible architectural guidance.

These technological developments represent ongoing challenges that will require continuous framework adaptation as the AI landscape evolves.

E. FUTURE RESEARCH DIRECTIONS

To address the identified limitations and enhance the PRISE framework's broader applicability, future research should prioritize the following areas:

Cross-Regional Validation: Systematic validation across different geographic regions and cultural contexts, particularly examining how cultural dimensions affect framework implementation and developing region-specific adaptation guidelines.

Industry Diversity Studies: Comprehensive validation across traditional industries including manufacturing, agriculture, and energy sectors, to clarify sector-specific adaptation requirements and regulatory constraints.

Organizational Scale Research: Examination of framework effectiveness across different organizational types, from startups to large enterprises, government agencies, and resource-constrained environments.

Technology Evolution Tracking: Investigation of how emerging AI technologies require framework modifications, developing technology-specific guidance for rapidly evolving AI capabilities.

Longitudinal Impact Studies: Long-term tracking of AI products developed using PRISE principles to understand sustained effectiveness and social implications.

Future work will include piloting the PRISE framework with external organizations in various industries and regions to further test and refine its generalizability. These efforts will help ensure the framework's adaptability and value beyond the Microsoft context.

These research directions will significantly strengthen the framework's theoretical foundation and practical utility across diverse AI innovation contexts.

APPENDIX
PRISE FRAMEWORK OPERATIONAL TOOLS

This appendix provides detailed templates and guidelines for the operational tools referenced throughout the PRISE framework. These tools serve as practical instruments for implementing each dimension effectively.

A. VALUE ASSESSMENT CANVAS

The Value Assessment Canvas structures value evaluation around individual pain points, with each canvas accommodating multiple related pain points. Table 9 presents the comprehensive template structure.

B. VALIDATION CARDS

Each Validation Card addresses a single hypothesis and contains the structured elements shown in Table 10.

C. VERTICAL SCENARIO EVALUATION TOOL

The Vertical Scenario Evaluation Tool provides a systematic approach to assess and prioritize different market opportunities. The scoring framework is detailed in Table 11.

The final score is calculated as a weighted sum: $\textit{Final Score} = 0.4 \times \textit{Value Potential} + 0.3 \times \textit{Technical Feasibility} + 0.2 \times \textit{Data Availability} + 0.1 \times \textit{Entry Barriers}$. Scenarios scoring above 3.5 are typically considered viable for PRISE framework implementation.

D. PRODUCTIZATION READINESS CHECKLIST

The Productization Readiness Checklist ensures AI products are prepared for production deployment across technical, operational, and business dimensions. Table 12 provides the comprehensive assessment framework.

Assessment Protocol: Each item is rated as Complete/In Progress/Not Started. Products should achieve "Complete" status for all items within their category before proceeding to production launch. Critical items (marked with *) require completion before any production deployment.

Usage Guidelines:

- Conduct initial assessment during Single-focus Development phase
- Review weekly during End-to-End Productization phase
- Require sign-off from technical, operational, and business stakeholders
- Use as gate criteria for production release decisions

REFERENCES

[1] S. Amershi, A. Begel, C. Bird, R. DeLine, H. Gall, E. Kamar, N. Nagappan, B. Nushi, and T. Zimmermann, "Software engineering for machine learning: A case study," in *Proc. IEEE/ACM 41st Int. Conf. Softw. Eng., Softw. Eng. Pract. (ICSE-SEIP)*, May 2019, pp. 291–300.

[2] C. Argyris and D. A. Schön, *Organizational Learning II: Theory, Method, and Practice*. Reading, MA, USA: Addison-Wesley, 1996.

[3] T. Brown, "Design thinking," *Harvard Bus. Rev.*, vol. 86, no. 6, p. 84, 2008.

[4] C. M. Christensen, *The Innovator's Dilemma: When New Technologies Cause Great Firms To Fail*. Brighton, MA, USA: Harvard Bus. Review Press, 1997.

[5] T. R. Eisenmann, E. Ries, and S. Dillard, "Hypothesis-driven entrepreneurship: The lean startup," Harvard Bus. school entrepreneurial Manage., Boston, MA, USA, Tech. Rep., 2012.

[6] *Gartner Predicts 30% Generative AI Projects Will Be Abandoned After Proof Concept By End 2025*, Gartner, Stamford, CT, USA.

[7] *Gartner Survey Finds Generative AI is Now the Most Frequently Deployed AI Solution in Organizations*, Gartner, Stamford, CT, USA.

[8] A. Griffin, "PDMA research on new product development practices: Updating trends and benchmarking best practices," *J. Product Innov. Manage.*, vol. 14, no. 6, pp. 429–458, Nov. 1997.

[9] D. Holliday, S. Wilson, and S. Stumpf, "Artificial intelligence and trust in human–computer interaction," in *Proc. CHI Conf. Extended Abstr. Human Factors Comput. Syst.*, 2017, pp. 2844–2851.

[10] *Overcoming GenAI Challenges: Key Insights for CDOs in 2025*, Inf., Redwood, CA, USA.

[11] A. Khurana and S. R. Rosenthal, "Towards holistic 'front ends' in new product development," *J. Product Innov. Manag., An Int. Publication Product Develop. Manage. Assoc.*, vol. 15, no. 1, pp. 57–74, 1998.

[12] D. Kreuzberger, N. Kühl, and S. Hirschl, "Machine learning operations (MLOps): Overview, definition, and architecture," *IEEE Access*, vol. 11, pp. 31866–31879, 2023.

[13] S. Lindberg, C. Rossitto, O. Knutsson, P. Karlström, and S. Männikkö Barbutiu, "Doing good business? Design leaders' perspectives on ethics in design," *Proc. ACM Hum.-Comput. Interact.*, vol. 8, pp. 1–22, Feb. 2024.

[14] L. E. Lwakatare, A. Raj, J. Bosch, H. H. Olsson, and I. Crnkovic, "A taxonomy of software engineering challenges for machine learning systems: An empirical investigation," in *Proc. 20th Int. Conf. Agile Processes Softw. Eng. Extreme Program.*, Montréal, QC, Canada. Cham, Switzerland: Springer, Jan. 2019, pp. 227–243.

[15] G. A. Moore, "Crossing the chasm: Marketing and selling technology products to mainstream customers," Harper Bus. Essentials, Broadway, NY, USA, Tech. Rep., 1991.

[16] L. M. Murphy and P. L. Edwards, "The valley of death and new public sector bridges: Bridging the innovation gap between applied and basic research," *J. Public Admin. Res. Theory*, vol. 16, no. 3, pp. 389–404, 2006.

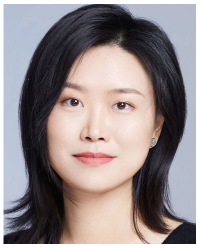
[17] A. Ng, "A chat with Andrew on MLOps: From model-centric to data-centric AI," Tech. Rep., 2021.

[18] I. Nonaka, "A dynamic theory of organizational knowledge creation," *Org. Sci.*, vol. 5, no. 1, pp. 14–37, Feb. 1994.
[19] in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2020.
[20] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 1135–1144.
[21] Everett M Rogers, *Diffusion of Innovations*. Mumbai, India: Free Press, 2003.
[22] N. Sambasivan, S. Kapania, H. Highfill, D. Akrong, P. Paritosh, and L. M. Aroyo, "Everyone wants to do the model work, not the data work: Data cascades in high-stakes AI," in *Proc. CHI Conf. Human Factors Comput. Syst.*, May 2021, pp. 1–15.
[23] D. Sculley, G. D. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young, J.-F. Crespo, and D. Dennison, "Hidden technical debt in machine learning systems," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, Dec. 2015, pp. 2503–2511.
[24] A. J. Maule and G. P. Hodgkinson, "Heuristics, biases and strategic decision making," *Current Directions Psychol. Sci.*, vol. 15, no. 2, pp. 68–71, Feb. 2002.
[25] J. Stewart, "Multiple-case study methods in governance-related research," *Public Manage. Rev.*, vol. 14, no. 1, pp. 67–82, Jan. 2012.
[26] L. G. Tornatzky, M. Fleischer, and A. K. Chakrabarti, *The Processes of Technological Innovation*. Lanham, MD, USA: Lexington books, 1990.
[27] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Quart.*, vol. 27, no. 3, pp. 425–478, 2003.
[28] R. K. Yin, *Case Study Research and Applications: Design and Methods*. Newbury Park, CA, USA: Sage, 2017.
[29] D. Zha, Z. P. Bhat, K.-H. Lai, F. Yang, and X. Hu, "Data-centric AI: Perspectives and challenges," in *Proc. SIAM Int. Conf. Data Mining (SDM)*, Jan. 2023, pp. 945–948.
[30] D. Zha, Z. P. Bhat, K.-H. Lai, F. Yang, Z. Jiang, S. Zhong, and X. Hu, "Data-centric artificial intelligence: A survey," *ACM Comput. Surveys*, vol. 57, no. 5, pp. 1–42, Jan. 2025.
[31] J. Zhang, R. V. Mummidi, and R. Panayappan, "Machine learning-based software application modernization assessments," U.S. Patent 11 354 120, Jun. 7, 2022.

BAOLI WANG received the B.S. degree from East China Jiaotong University, in 2014, and the M.S. and Ph.D. degrees from the Institute of Computing Technology, University of Chinese Academy of Sciences, Beijing, China, in 2022, with research focus on adversarial attack and defense in computer vision.

From 2020 to 2022, he was a Research Intern with the Face++ Megvii Technology Company Research Institute, Beijing. Since 2022, he has been a Software Development Engineer II with Microsoft Search Technology Center Asia (STCA), Beijing. His current research interests include operating system development, artificial intelligence coding, and an AI agent framework.

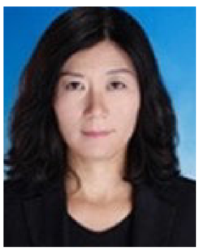
Dr. Wang received the Alibaba Security AI Challenger Award, in 2019, the First and Second Prizes in the Gitium Blockchain Hackathon, in 2018, and the Special Sponsor Award in the Wanchain Blockchain Hackathon, in 2018. He also ranked 43rd out of 1959 participants in the Alibaba Tianchi Big Data Competition, in 2017.



YE LI (Member, IEEE) was born in Beijing, China. She received the B.S. degree in information science from Peking University, Beijing, in 2000, and the M.S. degree in computer science and engineering from Beihang University, Beijing, in 2003.

She has more than 20 years of industrial experience. In 2010, she was a Senior Software Engineer with Sun Microsystems, where she helped develop the Sun Device Detection Tool and related certification suites. From 2010 to 2014, she was a Staff Software Engineer with Dell EMC, contributing to large-scale data analytics platforms. Since 2014, she has been with the Microsoft Search Technology Center Asia, where she is currently a Principal Applied Scientist. Her research interests include generative large language models, natural-language processing, knowledge graphs, and intelligent conversation systems. She has co-authored several conference papers on knowledge-graph construction and is a core contributor to the open-source platform SmartKG, in 2022, and GoTFlow, in 2024.

Ms. Li is a Senior Member of China Computer Federation (CCF). She has received multiple internal Microsoft awards, including the Global Hackathon Excellence Badge, in 2021 and 2023, and frequently serves as a Technical Mentor for industry AI-education initiatives.



WEN LEI was born in Wuhan, Hubei, China. She received the M.S. degree in computer science from the University of Pennsylvania, Philadelphia, PA, USA, in 1999.

She has more than 25 years of industrial experience. She began her career as a Software Engineer with Verizon Communications in the Greater Philadelphia Area, from 1999 to 2001. After returning to China, she held an Engineering Manager role at Sun Microsystems, until 2011. She then joined Dell EMC, as the Senior Engineering Director, in 2017. After that, she has been with the Microsoft Search Technology Center Asia, where she is currently a Partner Engineering Manager. Her work focuses on advertising platforms, marketplace services, and AI landing in enterprise software research and development.

Ms. Lei actively mentors early-career engineers and women engineers through her career and has received several awards for engineering leadership at both Dell EMC and Microsoft.