# RETHINKING REQUIREMENT ANALYSIS FOR AI-BASED PROJECTS

Sungho Kim [1], Sazid Rahman Kazi [2], Roise Uddin [3], Md Shahnawaj [2], and Yearanoor Khan [2]

[1] Department of Computer Science, Korea University, Seoul, Korea,
[2] Department of Information Technology, Pacific State University, Los Angeles, California, USA,
[3] Department of Master of businesses Administration-cybersecurity concentration, California State University, USA

***ABSTRACT***

*The rapid adoption of Artificial Intelligence (AI) across industries has revealed limitations in traditional requirement analysis methodologies, which were not designed to address the complexities and iterative nature of AI-based projects. This paper proposes a refined thought process for requirement analysis tailored to the needs of AI-driven initiatives, whether AI is the primary focus or an integrated component of a larger system. By emphasizing the dynamic interplay between data, models, and deployment environments, the proposed approach departs from linear methodologies, advocating for an adaptive and iterative process. Using case studies, we demonstrate how this concept ensures better alignment with business goals, enhances data utility, and improves model performance while addressing ethical considerations and practical constraints. This paper aims to provide practitioners, researchers, and project owners with actionable insights to optimize AI project outcomes in an increasingly complex technological landscape.*

***KEYWORDS***

*Requirement Analysis, AI Projects, Data-Centric AI, Model Selection, Deployment Strategy*

## 1. INTRODUCTION

Artificial Intelligence (AI) has emerged as a transformative technology, enabling innovative solutions across industries such as healthcare, finance, manufacturing, and entertainment. However, the development of AI-based systems introduces unique challenges that distinguish these projects from traditional software or system development. Central to these challenges is the requirement analysis phase—a critical step that lays the foundation for project success. Traditional methodologies often fall short in addressing the dynamic, data-driven, and iterative nature of AI projects [1], necessitating a re-evaluation of conventional approaches.

Unlike traditional projects, where requirements are typically static and well-defined, AI projects often encounter evolving requirements that change with data exploration, model iteration, and stakeholder insights [2]. Additionally, the success of AI systems depends not only on accurate model predictions but also on factors such as data quality, ethical considerations, deployment environments, and user trust. These complexities demand a requirement analysis framework tailored to the unique demands of AI projects.

David C. Wyld et al. (Eds): Signal, CDKP, AdNLP, MLT, SAI – 2025
pp. 147-158, 2025. CS & IT - CSCP 2025 DOI: 10.5121/csit.2025.150911

This paper explores a refined approach to requirement analysis for AI-based projects, proposing four distinct stages: Idea Generation and Validation, Research on Data, Model Selection, and Application Environment and Deployment. These stages represent the key pillars of requirement analysis, emphasizing iterative validation, dynamic requirement evolution, and the interplay between data, models, and deployment environments. Whether AI serves as the central focus or a supplementary component of a larger system, this methodology aims to provide a structured pathway to guide project teams and stakeholders through the intricacies of AI development.

## 2. Methodology

Our proposed methodology for requirement analysis in AI-based projects revolves around four interconnected stages. These stages are designed to address the iterative and dynamic nature of AI systems, ensuring alignment with project objectives and stakeholder expectations throughout the development lifecycle. The stages, described below, collectively form the backbone of requirement analysis for AI projects.

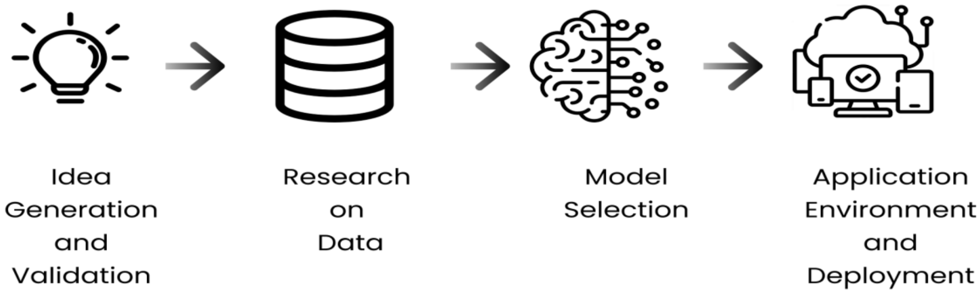


Figure 1. The Suggested Stages for Requirement Analysis

### 2.1. Idea Generation and Validation

The first stage involves defining the problem clearly and validating the feasibility of proposed solutions. This step lays the groundwork for all subsequent activities, focusing on establishing a shared understanding of the project's objectives, scope, and expected outcomes. Business analysts, data scientists, and subject-matter experts collaborate with stakeholders through structured workshops and focus groups to refine potential use cases and align the project's goals with business objectives.

Competitor and market analyses are integral to this stage, providing insights into existing solutions and industry trends. Feasibility assessments—spanning technical, operational, and financial dimensions—ensure that the project's foundation is sound. By the end of this stage, teams should have a validated problem statement and a clear roadmap for proceeding with requirement analysis.

### 2.2. Research on Data

Data forms the backbone of any AI project, making this stage critical to the overall process. At this point, the team identifies and evaluates potential data sources—ranging from internal databases and third-party providers to publicly available datasets—to determine whether they can adequately support the intended AI model. Rather than performing comprehensive data cleansing or transformation activities at this juncture, the primary focus is on assessing the availability, relevance, and sufficiency of data required to achieve the project's intended outcomes.

This phase also involves conducting a preliminary adequacy assessment, which may include reviewing data quality, completeness, and representativeness. Should the existing data prove insufficient or misaligned with the project's objectives, the team may need to reassess the project's scope or feasibility. By the conclusion of this stage, the team should possess a clear understanding of whether the existing data resources are suitable to justify proceeding with model experimentation and training.

### 2.3. Model Selection

Selecting the optimal model architecture is pivotal to the success of any AI project. During this stage, the team evaluates whether the dataset finalized in the preceding data research phase is suitable for fine-tuning an existing model. This involves benchmarking multiple algorithms against key performance indicators, such as accuracy, precision, recall, and F1 score, to determine both the feasibility and potential effectiveness of a fine-tuning approach.

If fine-tuning an existing model proves inadequate for meeting the project's objectives—whether due to performance gaps, evolving requirements, or data limitations—the team must consider developing a new model from the ground up. In this scenario, returning to the data research stage may be necessary to confirm that sufficient, high-quality datasets are readily available for comprehensive training.

To guide the decision-making process, a cost-benefit and time-to-delivery analysis is conducted, weighing the resource investment and expected timeframes of fine-tuning/developing an AI model. This ensures that the chosen strategy is both technically sound and aligned with overarching project constraints and objectives. Ultimately, the goal is to ensure the capability to produce an AI model that efficiently and accurately fulfills the intended requirements.

### 2.4. Application Environment and Deployment

This phase focuses on establishing a robust strategy for effectively deploying and operationalizing the project's final outcomes. Ensuring scalability, reliability, and maintainability is paramount, which involves selecting the appropriate infrastructure—whether cloud-based, on-premises, or a hybrid environment—to support the AI model under real-world conditions. In addition, security considerations and regulatory requirements may dictate the geographical location of deployment, particularly when handling sensitive data or adhering to compliance mandates.

For initiatives where the AI component integrates into a larger system, seamless alignment with existing or planned architectures is a key priority. By ensuring that the AI solution fits cohesively within the broader technology ecosystem, the project maximizes its overall value. This phase also entails evaluating maintenance and operational costs, providing a financial blueprint that helps guide decisions on infrastructure investments, support resources, and scaling strategies. Ultimately, the goal is to ensure the creation a sustainable, secure, and cost-effective deployment environment that fully supports the project's objectives while ensuring financial viability.

## 3. WHY THIS FRAMEWORK

Collectively, the four stages provide a structured framework to determine the technical, functional and financial feasibility of the AI components of a project. During the initial idea generation and validation phase, the teams clarify the objectives, refine use cases, and assess the overall viability of their proposed solutions. In the subsequent data-research stage, they verify that adequate, high-quality data are available, ensuring a solid foundation for model experimentation and training. The model selection process then allows teams to benchmark approaches, refine strategies, and ultimately choose whether to fine-tune an existing model or build a new one, guided by comprehensive cost-benefit analyses. Finally, careful planning of the application environment and the deployment strategy ensure that the AI solution can operate seamlessly, securely, and sustainably within the broader system architecture.

By going through these stages, organizations not only gain valuable insights into the real-world prospects of their project, but also acquire the clarity needed to make informed and evidence-based decisions. This thorough and methodical approach provides stakeholders with the confidence that the final solution will meet their intended goals while respecting operational constraints and efficient financial choice.

# 4. Literature Review

The rapidly evolving field of Artificial Intelligence (AI) has prompted researchers and practitioners to revisit and refine traditional requirement analysis methodologies. Although conventional approaches have proven effective for software and system development, they often fail to address the complexities inherent in AI-based projects. This section reviews the existing literature to identify gaps in current practices and establish the need for a redefined framework tailored to AI projects.

## 4.1. Challenges in Traditional Requirement Analysis

Traditional requirement analysis frameworks, such as Waterfall and Agile, operate effectively within the confines of well-defined and static requirements [3]. However, AI-based projects introduce an entirely new dimension of complexity. These projects rely on iterative processes, continuous model experimentation, and dynamic stakeholder expectations, which traditional methodologies struggle to accommodate.

Sommerville (2011) and Pressman (2020) underscore the rigidity of traditional requirement analysis approaches in software engineering, emphasizing upfront requirement definition as a cornerstone. This rigidity often results in project misalignment when applied to AI systems, where requirements evolve alongside data insights, model refinements, and changing business objectives. The Standish Group’s Chaos Report corroborates this by identifying requirement misalignment as a leading cause of project failure [2]. In AI contexts, this issue is magnified due to the dependency on high-quality data, nuanced ethical considerations, and model interpretability, which traditional methodologies fail to address comprehensively.

## 4.2. Emerging Approaches in AI Project Requirement Analysis

The limitations of conventional methods have catalyzed the development of adaptive and iterative frameworks tailored to AI projects. One notable trend is the integration of Agile principles with AI-specific feedback loops, enabling real-time adjustments based on ongoing discoveries in data and model performance. Proposals like an "AI-Agile" hybrid approach that emphasizes continuous validation and stakeholder collaboration, fostering an environment for dynamic requirement evolution .

Andrew Ng (2021) introduces the concept of data-centric AI, emphasizing the primacy of data quality and relevance over algorithmic optimization [4]. This perspective aligns with frameworks like CRISP-DM, which advocate iterative refinement and data-focused strategies.

### 4.3. Tools and Framework for AI-Specific Requirements

Various tools, such as Jupyter Notebooks, TensorFlow, and PyTorch, support exploratory data analysis and iterative model development [5]. Integrated platforms like Google Vertex AI and AWS SageMaker offer end-to-end AI lifecycle management but rarely provide structured guidance for the requirement analysis phase. While these tools streamline model development and deployment, they do not address the unique challenges posed by the interplay between data, models, and operational environments in AI projects.

Emerging frameworks for ethical AI, such as AI Explainability 360 and Aequitas, highlight the growing emphasis on fairness, accountability, and transparency in requirement analysis [6]. Despite these advancements, the integration of ethical considerations into a holistic requirement analysis framework remains underexplored.

### 4.4. Gaps in Current Literature

Existing literature underscores the necessity of adaptive and iterative methods but falls short in providing a unified framework for AI requirement analysis. Key gaps include:

- Evolving Requirements: Limited research addresses managing the dynamic interplay between data exploration, model iteration, and deployment environments, a hallmark of AI projects.
- Data-Driven Methodologies: While the importance of data quality is recognized, existing frameworks fail to provide actionable guidance on integrating data assessment into the requirement analysis lifecycle.
- Holistic Approaches: Most methodologies focus narrowly on specific project stages, such as model selection or data preprocessing, without addressing the interconnected nature of AI project components.

### 4.5. Conclusion of Literature Review

The review highlights the limitations of traditional requirement analysis methodologies in the context of AI projects. Emerging approaches and tools provide valuable starting points but lack the comprehensive scope needed to address the multifaceted nature of AI development. This paper aims to bridge these gaps by proposing a redefined framework that integrates the various stages, offering a robust foundation for successful AI project execution.

## 5. Hypothetical Case Studies

To illustrate the practical application of the proposed requirement analysis framework, this section presents detailed case studies. These examples showcase how the methodology can be adapted to address real-world challenges in diverse AI projects. Each case study highlights the stages of the methodology, including Idea Generation and Validation, Research on Data, Model Selection, and Application Environment and Deployment.

### 5.1. Case Study 1: AI Chatbot for Los Amigos Games Betting Association

#### 5.1.1. Objective

Develop a chatbot to provide sports-related data for informed betting decisions while maintaining ethical safeguards.

#### 5.1.2. Problem Statement

The client required a chatbot integrated into their website to allow subscribers to retrieve sports-related data, such as player statistics and historical performance. The system needed safeguards to prevent the chatbot from providing game predictions to mitigate misuse.

#### 5.1.3. Requirement Analysis

##### 5.1.3.1. Idea Generation and Validation

- Stakeholder workshops clarified the chatbot's primary functionality, ensuring alignment with business objectives and ethical boundaries
- Competitor analysis revealed gaps, such as the lack of advanced querying for sports statistics
- Feasibility assessments identified Retrieval-Augmented Generation (RAG) as an optimal model for the use case

##### 5.1.3.2. Research on Data

- The client’s existing sports database was assessed for sufficiency and relevance
- Preliminary data evaluation highlighted inconsistencies, leading to a structured cleaning strategy
- Metadata tagging, such as game locations and timeframes, for enhanced query relevance

##### 5.1.3.3. Model Selection

- Fine-tuning a RAG model was validated through benchmarking against diverse query types
- Ethical safeguards were integrated into the model training process, such as excluding predictive terms (‘predict,’ ‘forecast’)

##### 5.1.3.4. Application Environment and Deployment

- Early requirement analysis included evaluation deployment on the client’s web server infrastructure
- Security and scalability needs were outlined

#### 5.1.4. Outcome

The chatbot requirement ensured a robust and ethically compliant design that enhanced user engagement without misuse risks.

### 5.2. Case Study 2: AI-powered Vehicle Recommendation System

#### 5.2.1. Objective

Create a system to recommend vehicle options-buying, renting, or leasing-based on user preferences and budgets.

#### 5.2.2. Problem Statement

The client needed an AI system to analyze user budgets and preferences, recommend suitable options, and provide detailed insights on car types, production years, fuel efficiency, and mileage. The system also required integration of local rental and leasing data.

#### 5.2.3. Requirement Analysis

##### 5.2.3.1. Idea Generation and Validation

- Workshops with stakeholders identified a hybrid recommendation system as ideal, addressing the integration of purchase and rental options
- Feasibility assessments confirmed the viability of combining user preferences with real-time market data

##### 5.2.3.2. Research on Data

- Available datasets, including car-buying and rental database, were assessed for completeness and integration potential
- Missing rental, data prompted custom collection strategies and preprocessing for consistency
- Key attributes, such as fuel efficiency and mileage, were identified for recommendation optimization

##### 5.2.3.3. Model Selection

- Collaborative filtering algorithms were benchmarked for personalization
- Ethical considerations ensured unbiased recommendations across brands and pricing tiers

##### 5.2.3.4. Application Environment and Deployment

- The AI model needs to be integrated with a web-based application with a user-friendly interface
- Requirements analysis included infrastructure planning for scalability
- Real-time data input mechanisms were specified for continuous updates

#### 5.2.4. Outcome

The requirement analysis laid a strong foundation for developing a user-friendly, scalable recommendation system that maximized customer satisfaction

### 5.3. Case Study 3: Predictive Maintenance in Manufacturing

#### 5.3.1. Objective

Predict machinery failures and optimize maintenance schedules to minimize downtime and costs.

#### 5.3.2. Problem Statement

The client, a manufacturing firm, required a system to analyze machine sensor data and identify patterns indicative of potential failures. The solution needed to integrate seamlessly with existing production workflows.

#### 5.3.3. Requirement Analysis

##### 5.3.3.1. Idea Generation and Validation

- Stakeholder consultations defined predictive maintenance as a priority, aligning system objectives with operational needs
- Feasibility assessments focused on using existing sensor data to develop actionable insights

##### 5.3.3.2. Research on Data

- Historical sensor and maintenance data were evaluated for adequacy and quality
- Feature engineering identified critical indicators like vibrations levels and temperature thresholds

##### 5.3.3.3. Model Selection

- Time-series models, including Long Short-Term Memory (LSTM), were evaluated for their ability to handle sequential data
- Benchmarks ensured the model's reliability in diverse operational scenarios

##### 5.3.3.4. Application Environment and Deployment

- Requirement analysis specified edge computing for real-time analytics, minimizing cloud dependency
- Maintenance and monitoring protocols were outlined to support iterative improvements

#### 5.3.4. Outcome

The comprehensive requirement analysis ensured the feasibility and reliability of a predictive maintenance system tailored to the client's needs.

These case studies demonstrate the flexibility and effectiveness of the proposed requirement analysis framework in addressing diverse challenges across AI projects. By adapting the methodology to specific contexts, these projects achieved measurable improvements in efficiency, accuracy, and user satisfaction.

## 6. COMPARATIVE ANALYSIS

The proposed framework aims to redefine the requirement analysis stage for AI projects by ensuring that all key stakeholders—business analysts, subject matter experts, and decision-makers—gain a comprehensive understanding of the project lifecycle. This includes early insights into challenges, feasibility, cost matrix, timelines, and resource allocation. This section compares traditional, emerging, and proposed methodologies based on their ability to provide such holistic oversight during the requirement analysis phase.

## 6.1. Traditional Requirement Analysis Methodologies

### 6.1.1. Key Characteristics

Linear models like waterfall and iterative agile methodologies focus on defining static requirements at the outset of the project.

### 6.1.2. Strengths

- Efficient for projects with predictable and stable requirements
- Provides well-documented and straightforward workflows

### 6.1.3. Weaknesses

- Limited scope for stakeholder involvement beyond their specific areas of expertise
- Does not provide an integrated view of the project's feasibility, challenges, or resource demands
- Fails to address dynamic requirement changes, common in AI projects

### 6.1.4. Impact on Stakeholder Involvement

Stakeholders are often disconnected from the iterative processes, leading to gaps in understanding the end-to-end feasibility and risks of AI initiatives

## 6.2. Emerging Hybrid Frameworks

### 6.2.1. Key Characteristics

Modified agile or data-centric methodologies that adapt to the iterative nature of AI development.

### 6.2.2. Strengths

- Foster collaboration between data scientists and domain experts
- Integrate feedback loops that partially address evolving requirements
- Focus on data quality and iterative refinements of models

### 6.2.3. Weaknesses

- Offer fragmented insights into the project lifecycle, focusing more on technical adjustments than stakeholder engagement
- Ethical considerations and cost analyses are often deferred to later stages
- Lack a unified approach to evaluate project feasibility during the requirement analysis phase

### 6.2.4. Impact on Stakeholder Involvement

While stakeholder feedback is integrated intermittently, there is no structured mechanism to ensure a holistic understanding of the project's technical, financial, and operational dimensions.

## 6.3. Proposed Framework for AI Requirement Analysis

### 6.3.1. Key Characteristics

A four-stage framework—Idea Generation and Validation, Research on Data, Model Selection, and Application Environment and Deployment—that provides a comprehensive overview of the entire AI project during the requirement analysis stage.

### 6.3.2. Strengths

- Encourages extensive collaboration among stakeholders, including data analysts, subject matter experts, and business leaders, from the outset
- Ensures feasibility validation by providing early insights into technical challenges, cost implications, and timelines
- Enables stakeholders to foresee risks and constraints, aligning expectations with project goals
- Embeds ethical and compliance considerations within the requirement analysis phase, rather than deferring them to later stages

### 6.3.3. Weaknesses

- Demands initial investment in cross-functional collaboration and planning
- May require training or cultural shifts within teams unfamiliar with this integrated approach

### 6.3.4. Impact on Stakeholder Involvement

Stakeholders actively participate in shaping and validating the project from inception, gaining a clear picture of potential challenges, resource requirements, and operational constraints.

## 6.4. Comparative Insights

Table 1. Comparative insights on requirement analysis methodologies for AI projects

| **Feature** | **Traditional Approaches** | **Emerging Frameworks** | **Proposed Framework** |
|---|---|---|---|
| Stakeholder Engagement | Limited to specific phases | Moderate | Comprehensive and continuous |
| Feasibility Validation | Minimal | Partial | Integrated and holistic |
| Cost and Time Analysis | Addressed in later stages | Partially addressed | Core to the initial analysis |
| Dynamic Adaptability | Low | Medium | High |
| Ethical Considerations | Neglected | Secondary focus | Embedded in early stages |
| Project Lifecycle Oversight | Fragmented | Moderate | Complete |

### 6.5. Comparative Insights

The primary contribution of the proposed framework is its ability to empower stakeholders with a comprehensive understanding of the entire AI project lifecycle during the requirement analysis stage itself. Unlike traditional and emerging methodologies, this framework ensures that stakeholders can:

- Validate the feasibility of the project early on by examining challenges, costs, and timelines
- Align technical and business objectives through structured collaboration
- Anticipate risks and constraints, enabling more informed decision-making and resource allocation
- Embed ethical and operational considerations into the core planning stages, ensuring alignment with regulatory and societal expectations

By offering a holistic and inclusive approach to requirement analysis, the proposed framework not only addresses technical and operational needs but also bridges the gap between stakeholder expectations and project deliverables, making it a critical tool for the success of AI initiatives.

## 7. Comparative Analysis

The requirement analysis process for AI projects demands a unique approach distinct from traditional software development methodologies. This paper emphasizes that successful AI initiatives must begin with a robust assessment of project feasibility, encompassing idea validation, data availability, model suitability, and deployment strategies. Without these considerations, teams risk encountering significant challenges, such as poor data quality or inadequate model performance, which can derail the project and lead to stakeholder dissatisfaction.

By adopting the proposed framework, organizations can shift from superficial assessments to a thorough, AI-specific requirement analysis process. This ensures not only alignment between technical capabilities and business objectives but also proactive risk mitigation and informed decision-making. The framework's focus on iterative validation, stakeholder collaboration, and ethical considerations positions it as a critical tool for delivering AI projects that are scalable, reliable, and impactful.

Through this adaptive approach, stakeholders are empowered to navigate the complexities of AI development, bridging the gap between initial concepts and successful implementations. As AI continues to reshape industries, the methodologies outlined in this paper provide a comprehensive foundation to optimize outcomes and enhance the value of AI-driven solutions.

## Acknowledgments

Thanks, others, for any contributions, whether it be direct technical help or indirect assistance.

### References

[1] Roger S. Pressman, Bruce Maxim, "Software Engineering: A Practitioner's Approach," McGraw Hill, 2020

[2] The Standish Group, "Chaos Report," Standish Group International, 1995, https://www.csus.edu/indiv/v/velianitis/161/chaosreport.pdf

[3] Winston W. Royce, “Managing the Development of Large Software System,” Proceedings of IEEE WESCON, 1970

[4] Ng, A. MLOps: From Model-Centric to Data-Centric AI. 2021. Available online: https://www.youtube.com/watch?v=06-AZXmwHjo

[5] Adam Paszke et al., "PyTorch: An Imperative Style, High-Performance Deep Learning Library," Proceedings of the 33rd International Conference on Neural Information Processing Systems, 2019

[6] Pedro Saleiro et al., "Aequitas: A Bias and Fairness Audit Toolkit," arXiv preprint arXiv:1811.05577, 2018


## AUTHORS

**Sungho Kim** graduated from the Department of Computer Science at Korea University and is currently a professor at PSU California. His main research focus is on connecting automotive and medical technologies


**Sazid Rahman Kazi** is currently pursuing a Master of Science Information Systems concentration at Pacific State University ,LosAngeles, California, His primary research interests at the intersection of Information Technology, , Machine Learning, Deep Learning, Artificial Intelligence, He is focused on exploring how these domains can be integrated to develop innovative and secure intelligent systems.

**Roise Uddin**, graduated from the Department of Master of businesses Administration-cybersecurity concentration at California State University and is currently a IT manager at PSU California. His main research focus is on connecting Information TechnologycybersecurityMachine Learning Deep LearningAIComputer Vision

**Md Shahnawaj** is currently pursuing a Master of Science Information Systems concentration at Pacific State University, Los Angeles, California, His primary research interests at the intersection of Information Technology, Machine Learning, Deep Learning, Artificial Intelligence, and Computer Vision. He is focused on exploring how these domains can be integrated to develop innovative and secure intelligent systems.

**Yearanoor Khan** is currently pursuing a Master of Science Information Systems concentration at Pacific State University ,Los Angeles, California, His primary research interests at the intersection of Information Technology, , Machine Learning, Deep Learning, Artificial Intelligence