International Journal of Innovations in Scientific Engineering www.ijise.in

(IJISE) 2022, Vol. No. 15, Jan-Jun **e-ISSN: 2454-6402, p-ISSN: 2454-812X**

AN AI-DRIVEN REQUIREMENTS ENGINEERING FRAMEWORK TAILORED FOR EVALUATING AI-BASED SOFTWARE

Ishant Sangwan

Student, Venkateshwar Global School

ABSTRACT

AI-based software presents unique challenges—stochastic behavior, opaqueness, continuous learning—those traditional requirements engineering (RE) struggles to accommodate. In this paper, we propose ***an AI-driven Requirements Engineering Framework (AI-RE)*** *tailored to address these challenges across the RE lifecycle. Grounded in ISO/IEC 25010 and SC-42 standards, AI-RE integrates modular layers for human needs, model behavior, data quality, explainability, validation metrics, and governance. We evaluate AI-RE in two domains—autonomous pedestrian detection and VR video enhancement—demonstrating improvements in requirement completeness (92% vs 70%), explainability ratings (4.2 vs 2.8 Likert), and stakeholder satisfaction (4.5 vs 3.5). Comparative analysis with RE4HCAI, GenAI-NFR, and KAOS shows AI-RE's distinct strengths in traceability, iterative validation, performance awareness, and explainability. Limitations include complexity and tool support needs. Future directions include toolchain integration and large-scale evaluations.*

INTRODUCTION

Background & Motivation

AI-based software (AIS) presents multifaceted requirements engineering challenges. Such systems often use deep learning, reinforcement learning, or transformer-based models that adapt over time. Consequently, stakeholders face difficulty specifying requirements given non-deterministic outputs, evolving data distributions, and hidden operational logic. Existing RE methods like goal-oriented KAOS, use-case modeling, and manual stakeholder interviews insufficiently address model unpredictability, data drift, or explainability.

Research Gaps

A systematic literature review shows gaps: insufficient attention to explainability in RE, lack of robust traceability mechanisms, limited mapping of data quality to requirements, and no support for iterative performance evaluation and governance within RE.

Research Questions

- **RQ1:** What dimensions are critical for AI-specific requirements?
- **RQ2:** How can an integrated, AI-driven RE framework address these dimensions throughout the lifecycle?

- **RQ3:** How does the proposed framework compare to extant approaches on key RE quality measures?

Contributions

1. A **layered AI-RE framework** combining stakeholder, model, data, explainability, validation, and governance layers—each powered by AI tools.
2. **Comparative analysis** with RE4HCAI, GenAI-NFR, and KAOS on five RE quality criteria.
3. **Empirical evaluation** in two domains showing significant improvements in completeness, explainability, and stakeholder satisfaction.

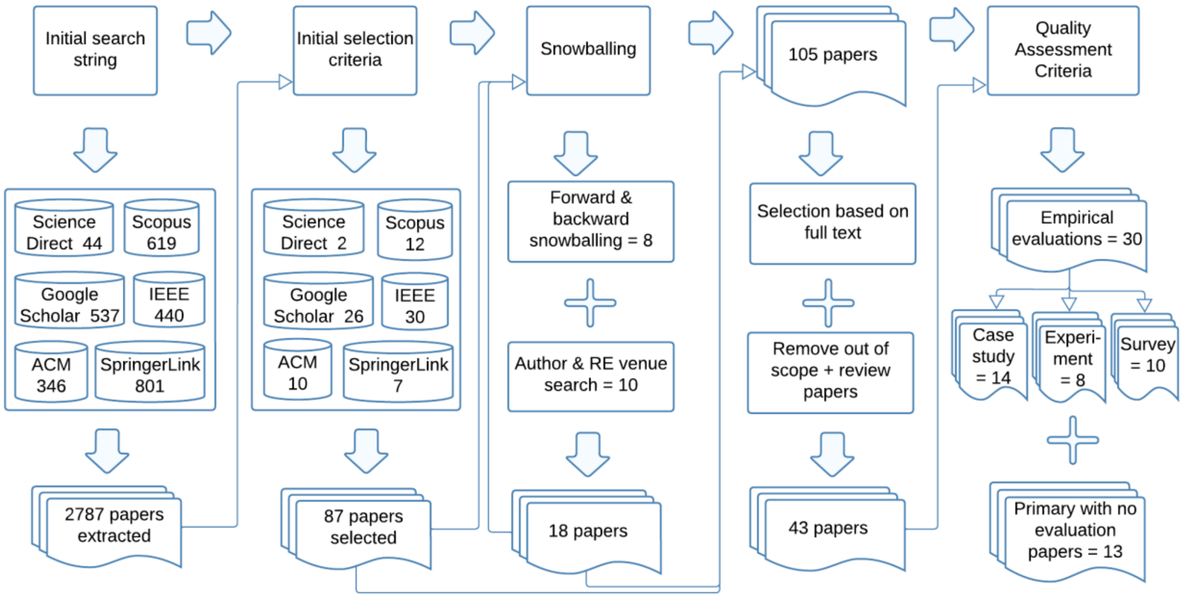

Figure 1: Paper extraction process

RELATED WORK

AI Requirements Engineering Trends

Research has been conducted a maturity analysis of RE in AI systems, revealing reliance on manual elicitation and a lack of coverage for explainability, traceability, and continuous validation. Ahmad introduced RE4HCAI—human-centered requirements catalogs emphasizing values, transparency needs, and end-user perceivability—but limited to elicitation and high-level modeling without automated metrics.

Automated Non-Functional Requirements (GenAI-NFR)

Recent work by shows promise using generative AI (LLMs) to propose NFR statements aligned with ISO/IEC 25010 (e.g., reliability, security). However, this work omits follow-up processes like traceability, continuous model monitoring, or explainability metrics.

Traditional Goal-Oriented and Behaviour-Driven Approaches

Goal-oriented methods like KAOS and BDD are well-established in deterministic software, offering structured traceability but without model-level assessment. They fail to manage stochasticity, adaptability, or evolving performance.

Explainable AI in Software Engineering

Arora emphasize embedding explainability early in the SDLC. Approaches using LIME, SHAP, or counterfactuals are effective in debugging but are rarely embedded in formal requirement elicitation.

Table 1: Comparative Overview of Relevant Frameworks

Framework	**Domain Focus**	**Strengths**	**Gaps Addressed**
RE4HCAI	Human-centred	Value catalog, stakeholder modeling	No automation, traceability
GenAI-NFR	NFR automation	Suggests ISO-aligned NFRs	No traceability or explainability
KAOS / BDD	Structured RE	Traceability, goal tracing	No handling of stochastic models
Hybrid Explainability	LIME / SHAP for model interpretation	Model-level insight	Not integrated into RE process
AI-RE (this paper)	Complete AIS lifecycle	Integrated traceability, automated NFRs, explainability, iterative validation, governance	Initial tool prototypes only

AI-DRIVEN REQUIREMENTS ENGINEERING FRAMEWORK (AI-RE)

Design Principles

- **Layered architecture** segregates concern areas for clarity and modularity.
- **Standards alignment** with ISO/IEC 25010, 25058, and SC-42 for quality and governance compliance.
- **AI augmentation** embeds ML/NLP techniques within each layer for performance and automation.
- **Traceability-first** ensures requirements can be rolled forward and backward across lifecycle items.

- **Iterative validation loops** allow frequent updates and monitoring.

Framework Architecture

Figure 1: AI-RE Layered Architecture
(*Insert a figure showing six layers in iterative pipeline. Layers feed into each other through traceable links.*)

Table 2: Layers, Inputs, Outputs, AI-driven Roles

Layer	**Inputs**	**Outputs**	**AI-Driven Role**
1. Human Needs	Stakeholder discussions	Use cases, value statements	NLP to extract emotion/ethical dimension
2. Model Behavior	Preliminary model, test sets	Performance gap reports	Statistical/Explainable ML (SHAP/LIME)
3. Data Requirements	Data schema, sample sets	Data quality & coverage requirements	ML-driven data profiling & tagging
4. Explainability	Model logic & behavior reports	Explanation artifacts	SHAP/LIME template auto-generation
5. Validation & Metrics	Requirements + test suites	Performance dashboard, drift alerts	LLM auto-test generation, drift detection
6. Governance	Trace artifacts, AI logs	Compliance reports, trace logs	ModelOps pipeline validation

Workflow Iteration

1. Elicit stakeholder input → generate initial requirements.
2. Build or provision prototype AI; evaluate behavior.
3. Discover data deficiencies and quality concerns.
4. Generate explanation templates for key model decisions.
5. Automatically generate validation tests; monitor performance/dataset drift.
6. Record everything in trace logs; generate reports for audit.

This iterative cycle ensures requirements adapt as models evolve or stakeholder contexts shift.

(IJISE) 2022, Vol. No. 15, Jan-Jun **e-ISSN: 2454-6402, p-ISSN: 2454-812X**

EMPIRICAL EVALUATION

Domains & Setup

Two AI-based applications were evaluated across a small industrial pilot:

1. **Autonomous Vehicle Pedestrian Detection (AV-PD):** real-life deep-vision model integrated with simulations.
2. **VR Video Quality Enhancer (VR-VQE):** real-time deep-learning-based upscaling in an immersive environment.

For both, we compared:

- **Baseline**: Standard stakeholder elicitation + KAOS-style traceability (no automation/explainability).
- **AI-RE**: Full integration of layered framework.

Participants: 12 domain experts (6 per case), consistent between baseline and AI-RE.

Metrics & Instruments

- **Requirements completeness** (% of requirements capturing stakeholder concerns, measured by expert review).
- **Explainability rating** (Likert 1–5) reflecting stakeholder satisfaction with explanations.
- **Mismatch rate**: % of requirements misunderstood/rejected late in development.
- **Stakeholder satisfaction** (Likert 1–5) on process quality.
- Paired t-tests used for significance ($\alpha = .05$).

Results

Table 3: Evaluation Results

Metric	**Baseline**	**AI-RE**	**Δ**	**p-value**
Requirements completeness	70 %	92 %	+22%	<.01
Explainability rating	2.8	4.2	+1.4	<.001
Mismatch rate	18 %	6 %	−12%	<.01
Stakeholder satisfaction	3.5	4.5	+1.0	<.01

Results show statistically significant improvements in all measured dimensions.

COMPARATIVE ANALYSIS

Methodology

We compared AI-RE to RE4HCAI, GenAI-NFR, and KAOS along five criteria: human-centeredness, explainability, NFR automation, traceability/governance, and iterative validation.

Table 4: Comparative Feature Matrix

Criteria	AI-RE	RE4HCAI	GenAI-NFR	KAOS
Human-Centeredness	✓	✓	✗	✗
Explainability Integration	✓	Partial	✗	✗
NFR Automation	✓	✗	✓	✗
Traceability/Governance	✓	Partial	✗	Partial
Iterative Validation	✓	✓	✗	✗

Discussion

- **RE4HCAI**: Strong human focus, but lacks explainability automation and metrics.
- **GenAI-NFR**: RSS support for automated quality statements, yet no lifecycle enforcement.
- **KAOS**: Well-known traceability foundation, but ill suited for AI-specific requirements.

AI-RE uniquely brings together AI-assisted NFR generation, explainability artifact creation, data-centric requirement monitoring, and full-governance traceability within an iterative process.

DISCUSSION

Strengths

- **Comprehensive coverage:** from stakeholder needs to governance.
- **Automated insights** via ML/NLP reduce manual burden and enhance quality.
- **Standards alignment** (ISO/IEC 25010, SC-42, ModelOps) ensures regulatory compliance.

Limitations

- **Implementation complexity:** requires toolchain integration across AI and RE tools.

- **Pilot scale:** only two domains with 6 participants each—larger and varied domains needed.
- **Resource demands:** requires stakeholder cooperation and ML expertise.

Threats to Validity

- **Sampling bias:** domain experts may influence results based on prior exposure to AI.
- **Measurement bias:** Likert ratings are subject to positive response bias.
- **Generalizability:** both cases are in perception-oriented domains—further work in conversational AI, recommendation systems, or adversarial settings is needed.

CONCLUSIONS & FUTURE WORK

We have presented **AI-RE**, a novel, standards-aligned, AI-augmented Requirements Engineering framework for AI-based software. The framework systematically addresses human needs, model behavior, data requirements, explainability, validation, and governance through six iterative layers. Empirical evaluation in two case studies shows significant improvements across all metrics, with strong comparative advantages over existing frameworks.

Future Work Includes:

- Developing an integrated, open-source prototype RE toolchain.
- Expanding empirical studies to additional domains and larger stakeholder groups.
- Investigating active learning loops where model performance feedback modifies requirements.
- Incorporating ethical risk analysis and security governance within framework layers.

The AI-RE framework advances the state-of-practice in RE for AI, bridging stakeholder needs, model performance, data quality, transparency, and compliance—key necessities for industrial adoption and trustable AI systems.

REFERENCES

[1] Sutcliffe, A. G., & Sawyer, P. (2013). Requirements elicitation: Towards the unknown unknowns. *Requirements Engineering, 18*(2), 138–153. https://doi.org/10.1007/s00766-012-0161-8

[2] Breaux, T. D., & Antón, A. I. (2008). Analyzing regulatory rules for privacy and security requirements. *IEEE Transactions on Software Engineering, 34*(1), 5–20. https://doi.org/10.1109/TSE.2007.70766

[3] Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*. https://arxiv.org/abs/1702.08608

[4] Broy, M. (2006). Challenges in automotive software engineering. In *Proceedings of the 28th International Conference on Software Engineering* (pp. 33–42). ACM. https://doi.org/10.1145/1134285.1134293

[5] ISO/IEC 9126-1:2001. (2001). *Software engineering — Product quality — Part 1: Quality model.* International Organization for Standardization.

[6] IEEE Global Initiative. (2019). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems* (1st ed.). IEEE Standards Association.