



The Ethical Requirements Stack: Operationalizing Adaptive Ethical Requirements with Human-AI Collaboration and GPT-Based LLMs

Mamia Agbese¹, Rebekah Rousi², and Pekka Abrahamsson³

¹ University of Jyväskylä, Seminaarinkatu, 40014 Jyväskylä, Finland
mamia.o.agbese@student.jyu.fi

² University of Vaasa, Wolffintie 32, 65200 Vaasa, Finland
rebekah.rousii@uwasa.fi

³ Tampere University, Kalevantie 4, 33100 Tampere, Finland
pekka.abrahamsson@tuni.fi

Abstract. The ongoing evolution and societal impact of AI systems demand systematic methods to embed ethics into AI development. Existing approaches often struggle to translate high-level ethical principles into concrete, adaptable software requirements, resulting in “ethical debt” that risks reputational harm, regulatory issues, and diminished stakeholder trust. This paper introduces the Ethical Requirements Stack (ERS), a structured, multi-layered artifact designed to elicit, decompose, and manage ethical requirements (ERs) from abstract themes to actionable development tasks. The ERS is operationalized through a human–AI collaborative workflow that leverages GPT-based Large Language Models (LLMs) for scalable ideation, complemented by human oversight to ensure contextual and ethical alignment. Using a design science research methodology, we demonstrate how the ERS supports the translation of stakeholder-elicited ethical values—aligned with frameworks such as IEEE 7000™-2021—into traceable software specifications. Our findings show that the ERS enables structured ethical reasoning and highlights the complementary strengths of AI-generated breadth and human critical judgment. This work contributes a practical approach for integrating ethics into the AI development lifecycle, supporting responsible innovation and reducing ethical debt through a combination of human-centered design and LLM-assisted requirements engineering.

Keywords: Ethical Requirements Stack · Human-AI Collaboration · Large Language Models (LLMs)

1 Introduction

The ongoing evolution of Artificial Intelligence (AI) continues to drive innovation in Software Engineering (SE). Modern AI-powered systems automate tasks, process vast datasets, generate content, act autonomously, and support

© The Author(s) 2026

G. Herzwurm et al. (Eds.): ICSOB 2025, LNBP 574, pp. 92–107, 2026.

https://doi.org/10.1007/978-3-032-14518-5_8

real-time interaction [1], transforming domains such as healthcare, finance, and autonomous systems [2–4]. While these advancements offer substantial benefits, they also introduce significant and evolving ethical challenges. Longstanding concerns, including bias, privacy, and accountability, are compounded by new risks such as AI-driven disinformation and unpredictable model behaviors [5]. Addressing these issues requires proactive integration of ethics across the AI development lifecycle to ensure sustainable and responsible software systems.

A central challenge lies in Requirements Engineering (RE): identifying, articulating, and managing ethical requirements (ERs)—explicit representations of values like fairness and transparency [6]. ERs are often abstract, interdisciplinary, and dynamic, evolving with rapid development cycles. Although Ethics by Design (EbD) approaches stress the importance of ERs, most frameworks lack adaptability for iterative environments, struggling to translate high-level principles into actionable and continuously evolving ERs [7, 8]. These shortcomings contribute to what can be described as *ethical debt*—a growing misalignment between high-level ethical principles and their concrete implementation in software systems. Similar to technical debt (the future cost of taking shortcuts), ethical debt accumulates when ethical considerations are deferred, only partially addressed, or inconsistently managed, creating risks that compound over time [9, 10]. Managing this debt requires mechanisms that make ERs explicit, traceable, and adaptable across the lifecycle, especially in iterative development contexts.

The IEEE Standard 7000™-2021 addresses this challenge by providing a process model for integrating ethical considerations into system design. The standard guides the elicitation of stakeholder values, their translation into ethical value requirements, and the traceability of these requirements from high-level principles to implementable specifications. It also emphasizes early integration, stakeholder engagement, and continuous alignment between ethical values and system functionality throughout the lifecycle [11]. Although the IEEE 7000 establishes a high-level process, practical instantiations that translate these principles into actionable methods, particularly in dynamic environments like Agile, remain limited.

To address this gap and extend the Ethical Requirements Stack (ERS) [12], this paper presents its design and operationalization within the Design and Demonstration phases of the Design Science Research Methodology (DSRM) [13]. The ERS provides a multi-layered structure that operationalizes ethical principles by translating them into context-sensitive and adaptable system requirements, aligning with standards such as IEEE 7000™-2021 and contributing to the mitigation of ethical debt. Designed for iterative methodologies like Agile, it provides a structured approach to integrating ethics into dynamic development processes. Our Demonstration validates the ERS’s core mechanisms—its multi-layered structure using human–AI collaboration—as a proof of concept. Central to this is the use of GPT-based large language models (LLMs), specifically ChatGPT, to enhance efficiency, while human oversight ensures contextual sensitivity and stakeholder awareness. This study examines how combining human and AI strengths can support the elicitation and refinement of ERs. Accordingly, we address the following research questions:

- **RQ1:** How can the ERS be operationalized to elicit, structure, and manage ethical requirements within an Agile environment?
- **RQ2:** What are the benefits and challenges of integrating Human–AI collaboration in the ERS workflow for generating and refining ethical requirements?

This study focuses on the design and demonstration of the ERS in early RE phases, situated within a broader Ethics by Design (EbD) for AI context. The primary contributions are threefold: (1) the design of the ERS as a structured, multi-layered artifact for managing adaptive ERs in RE; (2) The demonstration of Human–AI collaboration within the ERS workflow, showcasing the integration of GPT-based AI into requirements elicitation, refinement, and structuring; and (3) an empirical evaluation of the ERS in practice, assessing how Human–AI collaboration affects the comprehensiveness, clarity, and adaptability of ERs. The remainder of the paper is structured as follows: Sect. 2 reviews related work, Sect. 3 outlines the methodology, Sect. 4 presents the ERS artifact, the workflow, and findings, Sect. 5 discusses the evaluation, and concludes with limitations and future work.

2 Background and Related Work

Software Engineering (SE) faces new challenges as Artificial Intelligence (AI) systems continue to evolve and become integral to modern applications across various domains, including healthcare, finance, and autonomous systems [2–4]. AI, defined as computational systems capable of perceiving, learning, and making decisions with minimal human intervention [14], differs from traditional software that follows explicitly programmed rules, relying instead on data-driven models that introduce uncertainty, opacity, and adaptability [15]. These characteristics enable powerful capabilities in content generation, decision-making, and human–AI interaction [16, 17], but they also complicate established SE practices, particularly Requirements Engineering (RE). Traditional RE assumes that requirements can be elicited, specified, and validated, yet AI systems are adaptive, data-driven, and often opaque. This makes it difficult to translate stakeholder values into precise, testable requirements and to ensure traceability between requirements and system behavior [10].

The field of AI ethics addresses the moral, societal, and human-centered implications of AI technologies, emphasizing principles such as transparency, accountability, fairness, and privacy [18, 19]. However, rapid AI progress has introduced dynamic ethical challenges—disinformation, privacy breaches, and manipulative technologies like deepfakes and surveillance systems [3, 5]. Embedding ethics throughout the system lifecycle is essential. Yet, conventional AI ethics principles often remain abstract and difficult to operationalize, leading to implementation gaps and ethical blind spots [18]. These unresolved concerns accumulate as *ethical debt*: deferred or inconsistently managed ethical issues that, like technical debt, compound risks over time [9]. RE—the discipline of identifying, analyzing, and managing requirements—is a natural avenue for managing

this debt. Ethical Requirements (ERs) translate abstract principles into actionable system specifications [20], yet existing RE practices often struggle to capture and adapt ERs in fast-moving Agile contexts [10].

Several frameworks and methods have been developed to bridge this gap. IEEE Std 7000™-2021 provides a structured process for Value-Based Engineering (VbE), guiding stakeholder value elicitation, ethical value requirements (EVRs), and traceability from concept exploration to design [11]. While comprehensive, its phased, linear orientation can be misaligned with iterative Agile practices [7, 21]. ECCOLA offers a more Agile-oriented tool, organizing ethical themes into card-based prompts and practical recommendations [22]. It raises awareness but lacks mechanisms for systematically connecting high-level values to granular development tasks across iterations. Ethical User Stories (EUSs) adapt the user story format to embed ethical concerns (e.g., “As a user, I want my data anonymized so that my privacy is protected”) [23]. They are effective for capturing specific ethical needs but lack a hierarchical or evolving structure for managing ERs across project lifecycles.

These approaches have advanced the field but share common limitations: they are often static, difficult to adapt in dynamic AI contexts, and may overlook context-specific ethical risks [24]. As a result, ethical debt continues to accumulate, widening the gap between high-level principles and their actionable implementation. This creates space for frameworks that combine the rigor of standards like IEEE 7000 with the iterative flexibility of Agile methods. Human–AI (H-AI) collaboration provides a promising direction. Large Language Models (LLMs), such as GPT-based ChatGPT, can generate, organize, and refine ERs at scale, complementing human expertise in contextual understanding, moral reasoning, and oversight [25, 26]. While current research highlights the potential of H-AI collaboration in RE [27], there remains a lack of structured methodologies for systematically embedding it into the continuous elicitation and management of ERs.

The Ethical Requirements Stack (ERS) was first proposed in [12] as a multi-layered framework for translating abstract ethical values into Epics, Features, User Stories, and Tasks. Its conceptual strength lies in combining hierarchical structuring with iterative adaptability, aligning ethical requirements with Agile portfolio management principles. By explicitly linking ethical principles to evolving system requirements, the ERS offers a pathway for systematically managing and reducing ethical debt. This paper advances the ERS by presenting its detailed design and operationalization, focusing on how H-AI collaboration can enhance its workflow and address the limitations of existing frameworks.

3 Methodology

This study adopts an exploratory approach following the *Design Science Research (DSR)* paradigm to address the challenge of integrating ethical requirements (ERs) into AI development [28]. The Design Science Research Methodology (DSRM) by [13] is appropriate because the study involves both the creation

of an artifact—the *Ethical Requirements Stack (ERS)*—and the demonstration of its utility in a relevant setting. Accordingly, the research followed three phases: (1) designing the ERS as the core artifact, (2) demonstrating it through collaborative workshops with human–AI teams in an Agile-inspired environment, and (3) conducting a mixed-methods evaluation of its utility. Rather than testing hypotheses, the methodology emphasizes building and demonstrating a practical solution to a real-world problem [13, 28].

To complement DSR, we used the ECCOLA method [22, 29] and Ethical User Stories (EUSs) [23]. ECCOLA consists of 22 cards across nine AI ethics themes—analyze, transparency, data, agency and oversight, safety and security, fairness, wellbeing, accountability, and governance—each posing questions to prompt ethical reflection. EUSs adapt Agile user stories (“As a [user], I want [goal] so that [reason]”) to embed ethical reflections to help elicit ERs in AI development [23], where the “goal” states the ethical outcome and the “reason” its motivation. For example: “As a user, I want my data anonymized so that my privacy is protected.” This makes abstract ethical principles actionable, promoting ethically informed AI design [23].

3.1 ERS Design and Development (DSRM: Design)

The Ethical Requirements Stack (ERS), first introduced by [12], is a multi-layered framework for embedding ethical requirements (ERs) into AI development (see Fig. 1). The ERS operationalizes high-level ethical principles by systematically decomposing them into context-sensitive, actionable system requirements, maintaining traceability and alignment with standards and regulations, while supporting continuous feedback to adapt to evolving contexts. Inspired by Agile portfolio management and requirements stacks [30, 31], it organizes ERs hierarchically: *Themes* represent broad ethical principles; *Epics* define strategic objectives; *Features* describe necessary system functionalities; and *User Stories/Tasks* specify actionable development steps [32]. ERs are drawn from ethical guidelines, societal values, or AI ethics standards (e.g., IEEE 7000–2021), ranging from general mandates like “prevent discrimination” to specific ones such as “obtain informed consent before data collection.” This structure translates abstract principles into actionable requirements, promoting ethical alignment in practice. A core strength of the ERS is its adaptability, supporting ongoing refinement as new concerns or regulations arise. For this study, the ERS was refined through: (1) Conceptual adaptation of Agile principles; (2) Integration of ethical AI literature and tools like ECCOLA and Ethical User Stories (EUSs); and (3) Iterative structuring of abstract principles into practical, context-sensitive ERs [12].

3.2 ERS Operationalization: Workshops (DSRM: Demonstration)

The ERS was demonstrated through two scenario-based workshops (each approximately three hours), serving as exploratory case illustrations of its practical use. The workshops served as exploratory demonstrations of the ERS in practice

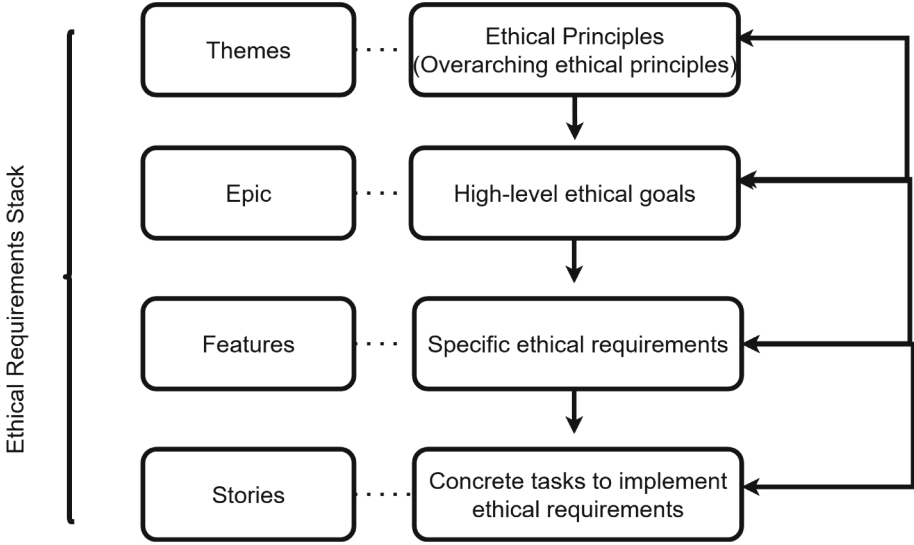


Fig. 1. Hierarchical organization of the ERS.

rather than as formal empirical validation. The outcomes of these workshops are reported in Sect. 4, where we illustrate how the ERS was operationalized in practice and what it yielded.

3.3 Participants and Data Collection

Ten researchers (Table 1) and ChatGPT collaborated in the study. Participants were purposively recruited from an international software engineering conference and a university research lab, with expertise in *Software Engineering, Information Systems, and Cybersecurity*—chosen for relevance to Agile methods, requirements engineering, sociotechnical insights, and security/privacy perspectives [33]. Their analytical skills and detailed feedback supported the exploration of the ERS’s practical use. All participants provided informed consent (redacted for anonymity [Consent Form](#)), which outlined anonymization measures, rights, voluntary participation, confidentiality, and potential risks. Workshop artifacts included human- and AI-generated EUSs and ERs, notes from ERS usability discussions, and Flamingo exports of collaborative work.

Table 1. Participant Roles and Expertise

Participant Role	Area of Expertise
R1–R2, R7–R9	Software Engineering
R3	Cyber Security
R4–R6, R10	Information Systems

Table 2. Central Ethical Themes and Selected ECCOLA Cards

ECCOLA Ethical Theme [18, 22, 29]	ECCOLA Ethical (Sub-theme) [22, 29]
Well-being	Societal Effect (#17)
Transparency	Communication (#3), Documenting Trade-offs (#4)
Data	Privacy and Data (#7)
Analyse	Stakeholder Analysis (#0)
Governance	Retention and Disposition (#21)
Agency/Oversight	Human Oversight (#11)

Agile-Inspired Setup. The workshops adopted Agile principles [34], structuring activities like sprints to enable iterative refinement. Cross-functional collaboration was simulated by having participants role-play managers, developers, and other stakeholders. Frequent feedback loops allowed continuous refinement of ERs for clarity, ethical alignment, and actionability [35].

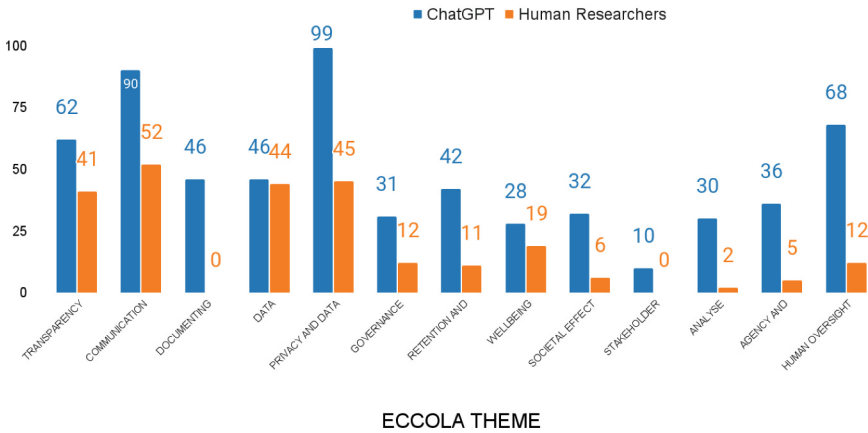
Scenario and Activities. The workshops focused on a fitness tracking app, chosen for its accessibility and relevant ethical challenges, such as privacy. The app enables users to monitor health data via smart devices (e.g., smartwatches, fitness bands, smartphones), requiring explicit permission for biometric data collection while also allowing manual input. A full use case description is provided in the supplementary materials at [Zenodo Repository](#). This scenario highlighted ethical concerns like privacy and transparency. Participants used seven ECCOLA cards spanning six ethical themes (Table 2) to develop Ethical User Stories (EUSs) and Ethical Requirements (ERs). Using the Flinga online whiteboard, they followed a four-stage process: (i) identifying ethical concerns via ECCOLA prompts, (ii) drafting EUSs, (iii) translating EUSs into system-level ERs, and (iv) refining ERs for clarity and alignment. For example, from the EUS “As a user, I want my fitness data to remain private...”, the ER “The system must enable users to delete their fitness data at any time” was derived by reframing the core ethical concern as a concrete, ECCOLA-aligned requirement. These ERs were then refined to enhance clarity, ethical fit, and actionability, enabling a collaborative, iterative approach to ethical design.

Human-AI Collaboration. ChatGPT was integrated to complement human input by generating ERs from three sources: the fitness app use case, ECCOLA cards (e.g., #7—Privacy and Data), and human-authored EUSs. For some themes (e.g., Stakeholder Analysis, Documenting Trade-offs), ERs were generated directly from the use case and ECCOLA prompts due to missing human-authored EUSs. Outputs were iteratively refined through prompt adjustments guiding ChatGPT toward ethical alignment and actionable formulation. Table 3 illustrates this process, showing how refinement shifted responses from vague user stories to operational requirements.

Table 3. Iterative Prompt Refinement During Human–AI Collaboration

Prompt Version	Description/Outcome
Initial Prompt	Example: “Using ECCOLA card #7—Privacy and Data, and this EUS, generate ERs.” Outcome: Responses included user stories and vague suggestions.
Refinement #1	Prompt: “Ensure ERs align with the ethical theme.” Outcome: Stronger alignment with ECCOLA themes.
Refinement #2	Prompt: “ERs should be actionable requirements, not user stories.” Outcome: Clarity improved; ERs became operationalizable

ECCOLA Theme Comparison - ChatGPT and Human Researchers

**Fig. 2.** Human vs AI ERs by ECCOLA theme.

3.4 Data Analysis

We employed a mixed-methods approach combining qualitative and quantitative analysis [36]. The ERS was managed in Microsoft Excel to structure ERs hierarchically (Theme → Epic → Feature → Story), remove duplicates and irrelevant items, and refine inconsistent phrasing. For **qualitative analysis**, Framework Analysis [37] was used to deductively code ERs against ECCOLA themes and map them into ERS layers by abstraction, with consensus checks ensuring reliability. For **quantitative analysis**, frequency counts of 869 ERs (620 from ChatGPT, 249 from humans) measured theme coverage and abstraction levels, highlighting differences in focus between human- and AI-generated ERs (Figs. 2, 3).

4 ERS as a Design Artifact

This section presents (i) the ERS as a design artifact, and (ii) outcomes from its operationalization in workshops.

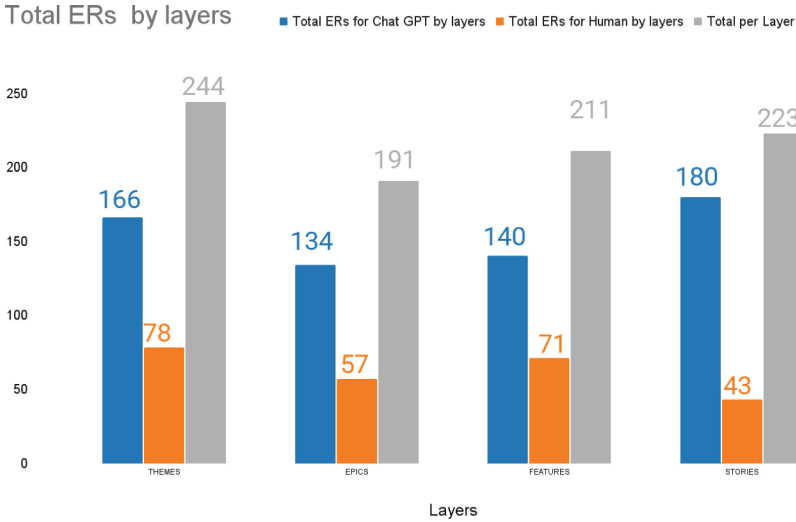


Fig. 3. Distribution of ERs across ERS layers.

The ERS was instantiated as a multi-layered artifact that aids the translation and management of abstract ethical principles to actionable requirements. Its hierarchy - (Themes, Epics, Features, Stories; Fig. 1) enables systematic decomposition of values elicited via tools like ECCOLA and EUSs. The design emphasizes three key value propositions for practitioners: *Traceability*, through hierarchical links from the Theme layer down to individual Stories; *Adaptability*, achieved through the granular, loosely coupled layers that allow requirement changes without system-wide refactoring; and *Scalability*, which supports the decomposition of broad ethical values across large, complex systems. By providing a formalized structure, the framework offers a means to address ethical debt as development contexts change. Table 4 shows how the privacy theme was decomposed during workshops, refining high-level principles (e.g., GDPR compliance) into implementable requirements.

Table 4. Illustrative ERS decomposition for Privacy

Layer	Example Requirement
Theme	Ensure GDPR compliance, including consent and deletion mechanisms
Epic	Users must be able to revoke data collection consent at any time
Feature	Provide user-friendly consent screens with permission controls
Story	Allow users to view, modify, or delete their data at any time

4.1 Results — What Operationalization Yielded

The workshops resulted in two instantiations of the ERS (available at [Zenodo Repository](#)). The first was a *human-only ERS*, derived entirely from participant-authored EUSs and ERs. The second was a *ChatGPT-assisted ERS*, where the AI generated ERs directly from human-authored EUSs based on ECCOLA prompts and the use case, with guided prompts to ensure contextual and ethical alignment. This hybrid instantiation broadened thematic coverage while preserving precision through human oversight. Together, these versions illustrate the ERS’s adaptability to different collaboration modes.

4.2 Findings

We present our findings aligned with the two research questions (RQs).

Answering RQ1: Design and Decomposition of the ERS. RQ1 examined how the ERS supports systematic decomposition and adaptive management of ethical requirements in Agile AI development. We assess its role across core Requirements Engineering (RE) stages: elicitation, analysis, specification, prioritization, validation, and management.

Elicitation drew on seven ECCOLA cards across six ethical themes, embedding ethics early in the process. *Analysis* involved clarifying ambiguities and checking feasibility against the fitness app scenario. *Specification* translated EUSs into preliminary ERs, refined into Themes, Epics, Features, and Stories (Table 3). *Prioritization* structured ERs into ERS layers by granularity and abstraction, ensuring traceability and prioritization [30]. *Validation and verification* were achieved through collaborative refinement for clarity and ethical alignment, though without deployment. *Management* positions the ERS as a dynamic backlog, enabling continuous adaptation across iterations. Table 4 illustrates the decomposition for Privacy-related requirements. Overall, the ERS operationalizes RE phases iteratively rather than sequentially, consistent with Agile principles.

Answering RQ2: Insights from Human-AI Collaboration. Our analysis of human- and AI-generated ERs produced five Preliminary Insights (PIs) on the benefits and challenges of Human-AI collaboration. The first three address quantitative traits; the last two focus on qualitative dynamics.

Quantitative Findings.

PI1: Coverage and Focus Divergence — ChatGPT generated a much higher volume of ERs with broad thematic coverage, while human researchers contributed fewer but more concentrated ERs.

A total of 869 ERs were produced (620 ChatGPT, 249 human). ChatGPT’s outputs spanned all ECCOLA themes, with a peak in *Privacy & Data*. In contrast, humans emphasized *Communication* while omitting others such as *Stake-*

holder Analysis. This divergence reflects the trade-off between AI’s breadth and humans’ selective depth.

PI2: Alignment with Core Themes — Both groups emphasized *Privacy & Data* and *Communication*, but *Stakeholder Analysis* was consistently underrepresented (Figure 2).

PI3: Layered Focus Divergence — ChatGPT concentrated on granular Story-level ERs, while humans focused on higher-level Themes and Epics, reflecting complementary strengths (Figure 3).

Qualitative Findings

PI4: Agile’s Iterative Refinement — Both ChatGPT and human researchers improved ERs through iterative cycles: prompt tuning for ChatGPT and group discussion for humans, enhancing clarity, alignment, and actionability.

Agile-inspired feedback loops supported continuous refinement. For ChatGPT, this involved adjusting prompts to yield clearer, more actionable ERs (Table 3), while for humans, collaborative discussion refined ERs for depth and contextual alignment. This iterative process produced structured requirements across ERS layers (Theme–Epic–Feature–Story), as illustrated in Table 4.

PI5: Synergistic Refinement — ChatGPT ensured breadth and scalability, while humans added contextual nuance and ethical precision, together producing layered, actionable ERs.

For instance, ChatGPT’s broad ERs (e.g., “protect user privacy”) became actionable only after refined prompting (Table 3). Human researchers further resolved ambiguities, grounding requirements in context and stakeholder needs. This synergy highlights a complementary model: AI accelerates ideation at scale, while human input ensures ethical clarity and implementation relevance.

5 Discussion and Conclusion

We synthesize key benefits and challenges, organizing the discussion thematically while explicitly referencing each Preliminary Insight (PI) to ensure empirical traceability.

5.1 The Benefits of Human-AI Synergy in ERS Development

The collaboration between human researchers and AI in the ERS process leverages their complementary strengths, enhancing scalability, actionability, and adaptability—without replacing human expertise.

Enhanced Comprehensiveness and Scalability (PI1): As identified in PI1, ChatGPT’s high-volume, cross-theme generation offers substantial scalability in early ethical analysis. In our study, ChatGPT generated 620 ERs compared to 249 from human researchers, ensuring broad initial coverage across ECCOLA themes. This reduces the risk of overlooking important ethical considerations, particularly in complex AI systems, and accelerates brainstorming in Agile environments [38]. By rapidly simulating stakeholder perspectives and mapping potential impacts, AI can enable teams to devote more human resources to refining and contextualizing ERs rather than starting from scratch [39].

Complementary Strengths and Actionability (PI3): PI3 highlighted a layered focus divergence that creates productive synergy. ChatGPT concentrated on story-level ERs—concrete, actionable items aligned with Agile’s specificity—while human researchers contributed more theme-level ERs, focusing on strategic and value-driven considerations. This combination bridges the gap between abstract ethical principles and operational system requirements. AI accelerates the production of actionable items, while human oversight ensures alignment with overarching ethical intent [40].

Dynamic and Adaptive Refinement (PI4, PI5): The iterative nature of Agile cycles, emphasized in PI4 and PI5, proved critical for adapting ethical frameworks in real time. Through successive feedback loops and iterative prompt adjustments, both AI- and human-generated ERs were refined for greater adaptability, precision, and contextual relevance. This continuous refinement transforms static frameworks into dynamic, evolving tools. AI’s speed enables quick updates in response to feedback, while human oversight ensures these adjustments remain ethically sound and contextually appropriate [41].

5.2 The Challenges of Human-AI Collaboration in ERS Development

Despite its benefits, our analysis uncovered persistent challenges that must be addressed to fully harness Human–AI synergy in RE.

Struggles with Contextual Nuance (PI2, PI5): As observed in PI2 and PI5, ChatGPT often lacked the contextual nuance needed for subtle ethical distinctions. For example, in PI2, the *Stakeholder Analysis* theme was underrepresented and occasionally misinterpreted, with the ECCOLA card treated as a principle rather than a contextual analysis tool [22]. This gap underscores AI’s limitations in translating framework guidance into real-world applications without targeted human intervention.

Risk of Blind Spots and Oversimplification (PI2, PI4): PI2 revealed a shared blind spot: both human and AI outputs underemphasized stakeholder inclusivity. Additionally, PI4 noted that iterative refinement—while beneficial—can lead to overfitting to a specific use case, risking oversimplification of broader ethical concepts and the erosion of nuance. Robust human oversight is therefore essential to prevent ethical debt and ensure generalizability [42].

Resource Strain from Oversight (PI1, PI5): While AI accelerates ER generation (PI1), the large volume of outputs requires careful human review to filter irrelevant or “hallucinated” requirements (PI5). This oversight involves active re-prompting, detailed contextual analysis, and strategic refinement, which can strain resources in development teams. Balancing AI-driven efficiency with the necessary depth of human oversight remains a practical challenge.

Implications for Agile Requirements Engineering: Our findings show a dynamic approach to embedding ethics in Agile RE by combining the *ERS framework* with *Human–AI collaboration*. The ERS decomposes high-level ethical requirements into actionable Stories linked to Epics, while human–AI workflows enable rapid ideation and nuanced refinement. This process translates abstract ethical values into concrete, adaptable requirements that evolve with technical development cycles [41].

Limitations: This study offers initial insights but has limitations. Internal validity is influenced by reliance on predefined themes and frequency analysis, which may miss nuance; qualitative coding and consensus discussions helped mitigate this. Researcher subjectivity and AI bias introduced variability, though transparency was supported by documenting researcher backgrounds. The single-use case limited diversity but allowed in-depth exploration, with findings aligned against established frameworks. External validity is constrained by the scope and use of a single ChatGPT version, though the interaction method is transferable. Limited Agile iterations restricted longitudinal observation, and the use of ECCOLA and prompting strategies shaped the results. Future work will explore multiple use cases, longitudinal settings, varied domains, and analysis of the ERs in the ERS [43–45].

5.3 Conclusion

This study applied a DSRM approach and Human–AI collaboration to generate the ERS for embedding ethics into Agile RE using a fitness app use case. Five key insights emerged: (1) AI generated broader, higher-volume ERs, while humans focused more deeply; (2) both emphasized *Privacy & Data* and *Communication*, but underrepresented *Stakeholder Analysis*; (3) AI leaned toward concrete Story-level ERs, humans toward abstract Themes; (4) Agile iterations supported ongoing refinement; and (5) combining AI scalability with human judgment improved clarity and nuance. Together, these findings show how the ERS dynamically operationalizes ethics—AI brings breadth, while humans provide ethical depth. This work sets the foundation for adapting the ERS across domains and Agile AI contexts.

Acknowledgments. I would like to acknowledge The Foundation for Economic Education (LSR) in Finland for its invaluable support.

References

1. OpenAI, R.: Gpt-4 technical report. arxiv 2303.08774. View in Article **2**(5) (2023)
2. Bhuyan, S.S., et al.: Generative artificial intelligence use in healthcare: opportunities for clinical excellence and administrative efficiency. *J. Med. Syst.* **49** (2025) Article 10
3. Reddy, S.: Generative ai in healthcare: an implementation science informed translational path on application, integration and governance. *Implementation Sci.* **19** (2024) Article 27
4. Jabbour, J., Reddi, V.J.: Generative ai agents in autonomous machines: a safety perspective. arXiv preprint [arXiv:2410.15489](https://arxiv.org/abs/2410.15489) (2024)
5. Alanazi, S., Asif, S.: Exploring deepfake technology: creation, consequences and countermeasures. *Hum.-Intell. Syst. Integr.* 1–12 (2024)
6. Habiba, U., Haug, M., Bogner, J., Wagner, S.: How mature is requirements engineering for ai-based systems? a systematic mapping study on practices, challenges, and future research directions. *Requirements Eng.* **29**, 567–600 (2024)
7. Cabrero-Daniel, B., Herda, T., Pichler, V., Eder, M.: Exploring human-ai collaboration in agile: customised llm meeting assistants. In: *International Conference on Agile Software Development*, Springer Nature Switzerland, Cham, pp. 163–178 (2024)
8. Fragiadakis, G., Diou, C., Kousiouris, G., Nikolaidou, M.: Evaluating human-ai collaboration: a review and methodological framework. arXiv preprint [arXiv:2407.19098](https://arxiv.org/abs/2407.19098) (2024)
9. Spiekermann, S., Winkler, T.: Ethical it innovation: a value-based system design approach. *Sci. Eng. Ethics* **21**(3), 527–545 (2015)
10. Vogel, C., Petersen, K.: Ethics in software engineering: a systematic literature review. *J. Syst. Softw.* **176**, 110925 (2021)
11. IEEE standards association: the standard model process for addressing ethical concerns during system design. Technical report (2021)
12. Agbese, M., Mohanani, R., Khan, A., Abrahamsson, P.: Implementing ai ethics: making sense of the ethical requirements. In: *Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering*, pp. 62–71 (2023)
13. Peffers, K., Rothenberger, M.A., Tuunanen, T., Chatterjee, S.: A design science research methodology for information systems research. In: *Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS)*, p. 219. IEEE (2007)
14. Leijnen, S., Aldewereld, H., van Belkom, R., Bijvank, R., Ossewaarde, R.: An agile framework for trustworthy ai. In: *NeHuAI@ ECAI*, pp. 75–78 (2020)
15. Alenezi, M., Akour, M.: Ai-driven innovations in software engineering: a review of current practices and future directions. *Appl. Sci.* **15**(3), 1344 (2025)
16. Müller, V.C.: Ethics of artificial intelligence and robotics. *Stanford Encyclopedia of Philosophy* (2020)
17. OpenAI: Chatgpt: optimizing language models for dialogue. <https://openai.com/chatgpt> (2023). Accessed 18 Feb 2024
18. Jobin, A., Ienca, M., Vayena, E.: The global landscape of ai ethics guidelines. *Nat. Mach. Intell.* **1**(9), 389–399 (2019)
19. Floridi, L.: *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford University Press (2023)

20. Liywalii, E.: Requirements engineering (re) in artificial intelligence (ai) systems implementation: the need to emphasize non-functional requirements (nfrs) for ethical ai. In Gerber, A., et al., eds.: SACAIR 2024, CCIS 2326, Springer, pp. 423–439 (2025)
21. Turk, D., France, R.: Assumptions underlying agile software development processes. *J. Softw. Syst. Model.* **17**, 19–25 (2003)
22. Vakkuri, V., Kemell, K.K., Jantunen, M., Halme, E., Abrahamsson, P.: Eccola—a method for implementing ethically aligned ai systems. *J. Syst. Softw.* **182**, 111067 (2021)
23. Halme, E., Jantunen, M., Vakkuri, V., Kemell, K.K., Abrahamsson, P.: Making ethics practical: User stories as a way of implementing ethical consideration in software engineering. *Inf. Softw. Technol.* **167**, 107379 (2024)
24. Lin, Z.: Beyond principlism: practical strategies for ethical ai use in research practices. *AI and Ethics*, pp. 1–12 (2024)
25. Song, B., Zhu, Q., Luo, J.: Human-ai collaboration by design. *Proc. Des. Soc.* **4**, 2247–2256 (2024)
26. Hasan, M.M., et al.: Human-ai collaboration in software design: a framework for efficient co-creation. *AIJMR-Adv. Int. J. Multidisciplinary Res.* **3**(1) (2025)
27. Gomez, C., Cho, S.M., Ke, S., Huang, C.M., Unberath, M.: Human-ai collaboration is not very collaborative yet: a taxonomy of interaction patterns in ai-assisted decision making from a systematic review. *Front. Comput. Sci.* **6**, 1521066 (2025)
28. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. *MIS Quart.* 75–105 (2004)
29. Agbese, M., et al.: Governance in ethical and trustworthy ai systems: extension of the eccola method for ai ethics governance using garp. *e-Inf. Softw. Eng. J.* **17**(1) (2023)
30. Cockburn, A., Highsmith, J.: Agile software development, the people factor. *Computer* **34**(11), 131–133 (2001)
31. Franch, X., Palomares, C., Quer, C., Chatzipetrou, P., Gorschek, T.: The state-of-practice in requirements specification: an extended interview study at 12 companies. *Requirements Eng.* **28**(3), 377–409 (2023)
32. Cohn, M.: User stories applied: for agile software development (2004)
33. Teddlie, C., Yu, F.: Mixed methods sampling: a typology with examples. *J. Mixed Methods Res.* **1**(1), 77–100 (2007)
34. Dingsøyr, T., Nerur, S., Balijepally, V., Moe, N.B.: A decade of agile methodologies: towards explaining agile software development (2012)
35. Thoring, K., Mueller, R., Badke-Schaub, P.: Workshops as a research method: guidelines for designing and evaluating artifacts through workshops. *Des. Sci.* **6** (2020)
36. Creswell, J.W., Clark, V.L.P.: *Designing and Conducting Mixed Methods Research*, 3rd edn. SAGE Publications, Thousand Oaks, CA (2018)
37. Ritchie, J., Lewis, J.: *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. SAGE Publications, London, UK (2003)
38. Resnik, D.B., Hosseini, M.: The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. *AI Ethics* 1–23 (2024)
39. Conti, C.J., Varde, A.S., Wang, W.: Human-robot collaboration with commonsense reasoning in smart manufacturing contexts. *IEEE Trans. Autom. Sci. Eng.* **19**(3), 1784–1797 (2022)
40. Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M., Floridi, L.: Operationalising ai ethics: barriers, enablers and next steps. *AI Soc.* 1–13 (2023)

41. Pflanzner, M., Traylor, Z., Lyons, J.B., Dubljević, V., Nam, C.S.: Ethics in human-ai teaming: principles and perspectives. *AI Ethics* **3**(3), 917–935 (2023)
42. Binns, R.: Fairness in machine learning: Lessons from political philosophy. In: *Conference on Fairness, Accountability and Transparency*, pp. 149–159. PMLR (2018)
43. Eswaran, U., Eswaran, V., Murali, K., Eswaran, V.: Human-centric ai balancing innovation with ethical considerations in the age of soft computing. In: *Soft Computing in Industry 5.0 for Sustainability*. Springer, Cham, pp. 87–116 (2024)
44. Odong, L.A., Perini, A., Susi, A.: Requirements engineering for collaborative artificial intelligence systems: a literature survey. In: *Research Challenges in Information Science*, Springer, Cham, pp. 409–425 (2022)
45. Taylor, C.S. In: *External Threats to Validity*. Oxford University Press, pp. 55–64 (2013)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

