

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Agbese, Mamia; Halme, Erika; Mohanani, Rahul; Abrahamsson, Pekka

Title: Towards a Business Case for AI Ethics

Year: 2024

Version: Published version

Copyright: © 2024 the Authors

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Agbese, M., Halme, E., Mohanani, R., & Abrahamsson, P. (2024). Towards a Business Case for AI Ethics. In S. Hyrynsalmi, J. Münch, K. Smolander, & J. Melegati (Eds.), *Software Business : 14th International Conference, ICSOB 2023, Lahti, Finland, November 27–29, 2023, Proceedings* (pp. 231-246). Springer. Lecture Notes in Business Information Processing.
https://doi.org/10.1007/978-3-031-53227-6_17



Towards a Business Case for AI Ethics

Mamia Agbese¹, Erika Halme¹, Rahul Mohanani¹,
and Pekka Abrahamsson²

¹ The University of Jyväskylä, Seminaarinkatu 15, 40014 Jyväskylä, Finland
{maoragbe,anermesi,rahul.p.mohanani}@jyu.fi
² Tampere University, Kalevantie 4, 33100 Tampere, Finland
pekka.abrahamsson@tuni.fi

Abstract. The increasing integration of artificial intelligence (AI) into software engineering (SE) highlights the need to prioritize ethical considerations within management practices. This study explores the effective identification, representation, and integration of ethical requirements guided by the principles of IEEE Std 7000–2021. Collaborating with 12 Finnish SE executives on an AI project in autonomous marine transport, we employed an ethical framework to generate 253 ethical user stories (EUS), prioritizing 177 across seven key requirements: traceability, communication, data quality, access to data, privacy and data, system security, and accessibility. We incorporate these requirements into a canvas model, the ethical requirements canvas. The canvas model serves as a practical business case tool in management practices. It not only facilitates the inclusion of ethical considerations but also highlights their business value, aiding management in understanding and discussing their significance in AI-enhanced environments.

Keywords: AI ethics · artificial intelligence · ethical requirements · IEEE Std 7000–2021 · ethical requirements canvas · software engineering

1 Introduction

The increasing integration of artificial intelligence (AI) into software engineering (SE) businesses is revolutionizing technology development, necessitating the incorporation of ethical requirements into management practices. This shift is emphasized by research [12, 30] and calls for aligning AI functionalities with ethical principles essential for guiding decision-making toward the development of trustworthy AI systems. Ethical requirements help to provide tangible actions derived from broader ethical principles like transparency, fairness, and privacy. For instance, the general principle of transparency becomes the need for “explainability” in AI, ensuring decision-making processes are clear and comprehensible for users [18]. As AI becomes more prevalent in sensitive sectors like health-care and education, SE organizations face increasing pressure from stakeholders, including developers, users, and regulators, to ensure AI systems like ChatGPT are not only innovative but also responsible and trustworthy [18, 30].

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBIP 500, pp. 231–246, 2024.

https://doi.org/10.1007/978-3-031-53227-6_17

Creating AI systems that are ethical and in sync with societal norms is a crucial aspect of trustworthy AI [12, 29]. Despite this, SE management stakeholders who guide decision-making find it challenging to incorporate ethical requirements into their practices effectively [1, 5, 12]. A primary challenge lies in these stakeholders' determination of ethical requirements relevant to business and representing them accordingly in their management approaches [1, 5]. This difficulty is compounded by a noticeable disconnect among these stakeholders in recognizing the value of ethical requirements [1, 5]. Existing ethical guidelines further exacerbate this gap, primarily focused on the technical aspects of SE projects, often neglecting the equally critical managerial dimensions that guide decision-making [25, 36]. This omission leads to the undervaluation of ethical considerations and puts organizations at risk of legal, reputational, and regulatory repercussions [1, 4].

To address the challenge faced by SE management stakeholders in determining and valuing ethical requirements in AI systems, our study utilizes the IEEE Standard Model Process for Addressing Ethical Concerns during System Design (IEEE Std 7000–2021) [19]. This standard serves as a vital tool for concept exploration and the development of the concept of operations (ConOps) stage, offering a comprehensive roadmap for embedding ethical considerations in the creation and operation of autonomous and intelligent systems (A/IS). It encourages managerial stakeholders to actively engage in four critical areas: *Identifying* relevant ethical requirements for their System of Interest (SOI), *Eliciting* these requirements based on applicability, *Prioritizing* their importance, and *Incorporating* them into management strategies, considering key stakeholder success factors. While the standard acknowledges that ethical consideration is not solely the responsibility of management, it underscores the pivotal role of management in establishing ethical benchmarks and supervising their outcomes. Consequently, our research is driven by two fundamental questions:

RQ1: *What ethical requirements do SE management stakeholders consider crucial for AI-empowered SOI?*; and **RQ2:** *How can ethical requirements be effectively evaluated and integrated as success factors in SE management strategies for AI-empowered SOI?*

The primary aim of this study is to underscore the crucial role of ethical requirements for SE businesses, particularly in AI-enhanced environments. By addressing the outlined research questions, we seek to guide organizations to circumvent ethical pitfalls and cultivate a culture of trustworthiness in AI development. Our objective is to contribute significantly to the ongoing conversation about integrating ethics into AI and SE practices, ultimately aiming to bolster stakeholder trust and position organizations as frontrunners in ethical AI deployment.

The remainder of this study is organized as follows: Sect. 2 provides an overview of the background and existing literature, while Sect. 3 describes our research methodology, including data collection, analysis, and key findings. Discussions based on our insights are presented in Sect. 4, and Sect. 5 offers the study's conclusions.

2 Background

AI ethics aims to ensure AI technologies are developed and utilized in alignment with ethical and societal values, preventing unforeseen consequences or damage. It examines the ethical principles and moral concerns tied to the creation, implementation, and usage of AI systems [26]. While AI ethics encompasses worries about machine behaviors and the potential emergence of singularity intelligent AI [26], this study doesn't explore that dimension. Issues like bias, surveillance, job displacement, transparency, safety, existential threats, and weaponized AI underscore the imperative of instilling ethical considerations into AI engineering. Consequently, private, public, and governmental stakeholders have set AI principles as ethical guidelines. Notable among these are the EU's trustworthy AI guidelines (AI HLEG), IEEE's Ethically Aligned Design (EAD), the Asilomar AI Principles, and the Montreal Declaration for Responsible AI [18,19]. Guiding principles distilled from various guidelines, as outlined by Ryan and Stahl [32] and Jobin et al. [21], include Transparency, Justice, Non-maleficence, Responsibility, Privacy, Beneficence, Autonomy, Trust, Sustainability, Dignity, and Solidarity.

2.1 Ethical Requirements

Ethical requirements are multifaceted, requiring careful consideration and interdisciplinary collaboration spanning technology, law, philosophy, and social sciences [24]. Ethical requirements of AI are primarily from foundational ethical principles or rules, such as transparency and fairness, and are pivotal for fostering trustworthy AI [15]. They help interpret the guiding principles and standards that ensure AI systems' ethical design, creation, deployment, and operation. From the principle of privacy, for instance, an ethical requirement is privacy and data protection, entailing that AI systems should handle personal and sensitive data carefully according to legal regulations and best practices [15,21]. As such, they help build trust and align AI endeavors with human values and societal aspirations [15]. However, in SE, ethical requirements are predominantly articulated as functional and non-functional requirements during the development phase [15], yet they are seldom addressed at the management level, typically only insofar as to meet legal mandates like the General Data Protection Regulation (GDPR) [1,24].

2.2 Trustworthy AI

With the increasing integration of AI across various aspects of human life, the concept of Trustworthy AI has evolved to encompass a broader range of societal and environmental considerations. These include the implications for employment, societal equity, and the environment. Despite the presence of specific frameworks and guidelines from organizations, governments, and international bodies, the critical requirements that truly define what makes AI trustworthy remain a central concern [12,29]. The AI HLEG and IEEE EAD have been

instrumental in identifying critical ethical requirements, significantly shaping the discourse on trustworthy AI [18,19]. These frameworks outline key ethical principles that serve as a guide for both academia and industry professionals. The AI HLEG highlights seven key requirements for trustworthy AI: human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, and accountability. Concurrently, the IEEE EAD emphasizes five: human rights, well-being, accountability, transparency, and awareness of AI's potential for misuse [19]. There's notable convergence in these requirements, which we explain as follows: *Human agency and oversight*: Emphasizes the importance of human rights and underscores the indispensability of human direction and supervision. *Technical robustness and safety*: Stresses the importance of crafting AI systems that resist threats, prioritize safety, have inherent protective mechanisms, and exhibit consistent, dependable, and replicable outcomes. *Privacy and Data Governance*: Navigates the privacy terrain, advocating the cause of data integrity, quality, and accessibility. *Transparency*: Entails a commitment to traceability, explainability, and effective communication of AI processes. *Diversity, non-discrimination, and fairness*: Encourages equitable AI practices, advocating for unbiased algorithms, universal design principles, and inclusive stakeholder engagement. *Societal and environmental well-being*: Focuses on AI's societal imprint, ranging from its ecological footprint to its broader societal repercussions and democratic implications. *Accountability*: Encompasses regularized auditing, transparent reporting, harm minimization, and effective remedial mechanisms. These enumerated requirements find application in tools like ECCOLA and Ethical User Stories (EUS), pivotal in executing the IEEE Std 7000–2021 approach of this study.

ECCOLA is an Agile-oriented method designed to enhance awareness and execution of AI ethics for developers in SE [36]. It synthesizes ethical requirements from AI HLEG and EAD, consolidating them into seven core themes or requirements and sub-requirements. The ECCOLA approach is a 21-card deck organized around seven primary requirements: transparency, data agency and oversight, safety and security, fairness, well-being, and accountability, and a stakeholder analysis card. Each requirement is represented further by one to six dedicated sub-requirement cards. ECCOLA is segmented into three components: the rationale behind its importance, actionable recommendations, and a tangible real-world example [36]. For direct access to ECCOLA, click here.

Ethical User Story concept integrates the user story methodology with an ethical toolset, facilitating the extraction of ethical requirements during technological design or development processes [16]. In SE and Agile methodologies, user stories help bridge business objectives and development activities by succinctly capturing customer demands [10]. These stories act as conduits to foster understanding between developers and users. They distill intricate concepts into more targeted information pieces, bolstering communication and collaboration

to ensure goal alignment. A standard user story is structured as: “As a [user role], I want [goal or need] so that [reason or benefit].” Here, the “user role” delineates a specific user’s identity or function. The “goal or need” specifies the desired outcome from the software, while the “reason or benefit” pinpoints the underlying motivation or value that drives this desire helping to concisely and clearly describe a user’s requirement for the SOI [10].

2.3 Standard Model Process for Addressing Ethical Concerns During System Design

The IEEE Std 7000–2021 provides a practical approach for SE businesses to identify and address ethical issues during the system design of their system of interest (SOI). We focus on the concept exploration and development of the concept of operations (ConOps) stage in our study, which emphasizes proactive communication with stakeholders, to help identify and prioritize ethical values to be integrated at the system design stage [20]. The procedure entails discerning these values from the operational concept, which lays out the system’s functionality, and from the value propositions and dispositions, which highlight the system’s benefits and potential outcomes. Central to the IEEE Std 7000–2021 are the Ethical Value Requirements (EVRs) concept. EVRs epitomize the essential worth of ethical requirements, ensuring that systems resonate with societal standards and uphold human rights, dignity, and well-being [12, 18, 20]. The standard advocates for meaningful engagement of primary stakeholders, especially those in management roles, throughout the design phase in *Identifying* pertinent ethical requirements by scrutinizing relevant ethical regulations, policies, and guidelines, including gathering stakeholder feedback. - *Eliciting* these ethical requirements based on their relevance to the SOI. - *Prioritizing* the inherent value of these requirements. - *incorporating* these values into the system’s core objectives and ensuring consistent communication and compliance monitoring with all concerned parties. Defining and embedding ethical requirements can bolster SOIs’ credibility, trustworthiness, and perceived value to help weave them seamlessly into their system’s design and development [20].

2.4 Implementing Ethical Requirements in SE Management

Aligning software development with an organization’s objectives is primarily achieved through SE management, which integrates critical success factors into operational and decision-making frameworks [14, 28]. Despite its importance, there’s a scarcity of tools that embed ethical requirements within SE management [3, 5]. Notably, the adaptation of canvas models for ethical representation is gaining traction among researchers and practitioners seeking to elevate ethical considerations in their practices [22, 27, 37]. Canvas tools are graphical representations that clarify intricate business concepts, facilitating stakeholder alignment. They break down various business facets, like customer segments or value propositions, into an easily digestible format often serving as a business snapshot enhancing understanding and communication [8, 28]. Some notable approaches

for the canvas model include The Ethics Canvas [22] which leverages the foundational blocks of the business model canvas to stimulate discussions on the ethical implications of technology. However, its scope on ethics is extensive and doesn't precisely target AI ethics or its requirements. The Open Data Institute's Data Ethics Canvas [27] offers a lens through which data practices can be ethically evaluated. Vidgen et al. [37] introduce a business ethics canvas, drawing inspiration from the applied ethics principles of the Markkula Center, which focuses on addressing data-centric ethical issues in business analytics. The canvas, however, predominantly focuses on the data ethics dimension. A more comprehensive canvas approach is the Trustworthy AI Implementation (TAII) canvas [2], which extends from the TAI framework [3]. It outlines the interplay of ethics within a company's broader ecosystem, touching upon corporate values, business strategies, and overarching principles but does not precisely pinpoint ethical requirements, potentially making it challenging for SE management stakeholders to translate it into actionable management practices [3].

3 Research Methodology

We adopt an exploratory approach to address our research questions. This approach is in line with Hevner et al.'s Design Science method, particularly the "build" component, given the innovative nature of our study and the limited resources in existing literature [17]. Exploratory methods provide valuable flexibility, especially when delving into less-explored research areas [35]. Hevner et al. emphasize the importance of adapting their seven guidelines, and our primary focus lies in developing conceptual artifacts, as outlined in their "Design as an artifact" guideline. While this phase typically yields conceptual insights rather than fully developed systems, the design science approach is crucial for shaping novel artifacts, even in the face of challenges [17].

3.1 Data Collection

We collaborated with 12 Finnish SE executives on an AI-enhanced project focused on autonomous marine transport for emission reduction and the enhancement of passenger and cargo experiences at the concept exploration stage. These executives represent various businesses specializing in different aspects of intelligent and autonomous SE, as detailed in Table 1. Our objective was to identify the essential ethical requirements these stakeholders deemed necessary for the AI-enabled System of Interest (SOI). To initiate our study, we secured the informed consent of our industry partners, emphasizing their entitlement to withdraw or request data deletion at any phase. Leveraging their SE background, which granted them a foundational understanding of the concepts, we embarked on a collaborative project segmented into three specific use cases. A series of workshops grounded on the brainstorming technique delineated by [33] facilitated the familiarization process with critical frameworks, including IEEE Std 7000-2021, ECCOLA, and the EUS concept.

During these sessions, the participants, who were predominantly executives, actively engaged in selecting pertinent ethical requirements from the 21 ECCOLA cards, highlighting those that resonated significantly with their business operations. The focus coalesced around ethical themes encapsulated by cards # 2 Explainability, # 3 Communication, # 5 Traceability, # 7 Privacy and Data, # 8 Data Quality, # 9 Access to Data, #12 System Security, # 13 System Safety, # 14 Accessibility, # 16 Environmental Impact, and # 18 Auditability. This careful selection served as a guide to pinpoint the ethical themes critical to their enterprise, facilitating a nuanced exploration. Extensive notes were documented to address subsequent inquiries and emerging concerns.

Table 1. SE Management Stakeholders

Solution provider	Area of expertise
Solution provider 1	Animation
Solution provider 2	Software development
Solution provider 3	Intelligent logistics
Solution provider 4	Remote and autonomous solutions
Solution provider 5	Transportation logistics
Solution provider 6	Computer controlled machinery
Solution provider 7	Intelligent translations
Solution provider 8	Intelligent transport infrastructure and logistics
Solution provider 9	Intelligent logistics
Solution provider 10	Information solutions
Solution provider 11	Automation solutions
Solution provider 12	Intelligent logistics

In eight workshops, each spanning one to three hours, we collaboratively formulated EUS using the ECCOLA method, tailoring the selections from ECCOLA to suit the requirements of each specific use case. Our detailed notes amounted to a total of 367, resulting in the creation of 253 EUS instances [34]. Examples of these instances include:

“As a [company CEO], with automated truck deliveries, I want [to have information, before sending my trucks on how data is handled], so that [I can feel secure that my data will not leak to unwanted parties].”

“As a [company data protection manager], I want to [authenticate the collected data] so that I can [ensure validity].”

“As a [system administrator], I want to [streamline the management of GDPR requirements] so that I can [ensure that the service remains unaffected by user information or data erasure requests].”

“As a [project stakeholder], I want the system [to feature clear and explainable logic] to [prevent project overruns or operational errors caused by unclear system descriptions].”

3.2 Data Analysis

We conducted our analysis utilizing content analysis, a systematic approach for dissecting qualitative data to discern recurring themes, patterns, and categories, ultimately yielding valuable insights [39]. In analyzing the EUS, we adopted an interpretive content analysis approach, prioritizing narrative interpretations of meaning over purely statistical inferences. This method enabled us to differentiate between manifest content, which represents overt messages in communication, and latent content, which encompasses subtle or underlying implications [39]. To streamline the analysis, we established a coding system. For instance, ‘TR’ was used as a code to symbolize ‘transparency’, while ‘DA’ represented ‘data’. These are just some examples of the various codes we employed throughout our analysis. These codes were then used to highlight specific ethical requirements within the dataset. For example, ‘TR’ pinpointed instances where transparency was a focal point in user stories. As we observed emerging patterns, we sought to identify correlations between the codes and overarching themes. These themes were then cross-referenced with central themes from the ECCOLA cards.

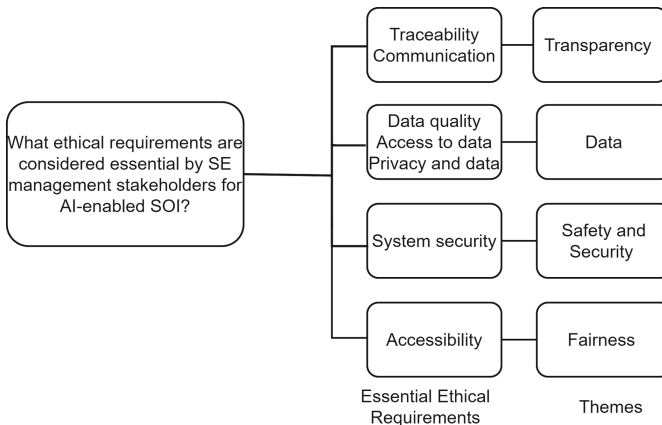
Utilizing the MoSCoW Prioritization technique [11], a popular tool in project management, software development, and business analysis, the executives classified the EUS based on their significance of “Must have, Should have, Could have, and Won’t have”. “Must have” captures indispensable requirements without which the project is incomplete. “Should have” comprised valuable yet non-critical elements; their omission wouldn’t jeopardize the project. “Could have” entails requirements that, while beneficial, aren’t urgent and can be tackled if resources permit. “Won’t have” covers those that are either irrelevant to the current project or simply unfeasible, possibly deferring them for later consideration or omitting them altogether [11]. The comprehensive prioritization can be found in Table 2. Of the 12 industry partners, nine participated in these classification exercises, while three were unavailable (denoted as N/A). The activity spanned several sessions, resulting in 177 out of the 253 EUS receiving priority rankings.

3.3 Findings

The prioritization from the EUS yielded seven distinct sub-requirements, categorized under four primary requirements. These sub-requirements are #5 Traceability, #3 Communication, #8 Data quality, #9 Access to data, #7 Privacy and data, #12 System security, and #14 Accessibility. They fall under the broader categories of Transparency, Data, Safety and Security, and Fairness. These emerged as crucial for SE management stakeholders, as illustrated in Fig. 1.

Table 2. Prioritization breakdown

Solution provider	Themes	Sub-Requirement	Prioritization
Solution provider 7	Transparency	#3 Communication	18
Solution provider 2	Transparency	#5 Traceability	42
Solution provider 10	Data	#7 Privacy and Data	20
Solution provider 9	Data	#7 Privacy and Data	29
Solution provider 12	Data	#8 Data Quality	7
Solution provider 4	Data	#8 Data Quality	7
Solution provider 6	Data	#9 Access to Data	3
Solution provider 2	Data	#9 Access to Data	5
Solution provider 10	Data	#9 Access to Data	4
Solution provider 11	Data	#9 Access to Data	3
Solution provider 12	Data	#9 Access to Data	3
Solution provider 8	Data	#9 Access to Data	2
Solution provider 9	Safety & Security	#12 System Security	24
Solution provider 7	Fairness	#14 Accessibility	10
Solution provider 1	N/A	N/A	-
Solution provider 3	N/A	N/A	-
Solution provider 5	N/A	N/A	-
Sum total of EUS	-	-	177

**Fig. 1.** Essential Ethical Requirements

4 Discussion

We examine our findings within existing research.

4.1 Essential Ethical Requirements

We analyze the seven identified ethical requirements and explore their significance and implications for stakeholders in SE management.

Traceability is pivotal in enhancing transparency and ensuring accountability within AI systems. It provides stakeholders with vital information to scrutinize and interpret the system's decisions [36]. By prioritizing traceability, those in SE management roles can effectively identify and manage the inherent risks associated with AI technology. This focus requires a detailed documentation process encompassing data sources, applied algorithms, computational models, and justifying particular outputs. Such comprehensive records identify potential weak points that could be prone to errors or biases, thereby enabling risk mitigation strategies to be deployed proactively [21]. As Ryan et al. underscore [32], maintaining stringent traceability practices reinforces accountability and fortifies customer and stakeholder trust, consequently elevating the organization's reputation.

Communication is central to disseminating essential details about an AI system's architecture, development phases, and functionalities to all pertinent stakeholders. Effective communication involves transparently articulating the system's objectives, capabilities, limitations, and possible repercussions. By doing so, stakeholders engaged in the project can gain a well-rounded understanding of the initiative's scope and aims, allowing them to identify and proactively address technical and ethical challenges. Open and transparent dialogue among SE management stakeholders can facilitate collaborative problem-solving and mitigate potential adverse outcomes. One challenge in communication within SE management is the complexity of technical jargon and the volume of information related to AI project documentation. However, prioritizing strategic communication can align expectations and clarify objectives [32].

Data Quality ensures that data serves its designated purpose and can be relied upon for making well-informed decisions within AI systems [6, 18, 23]. For SE management, data quality is a strategic component that influences the efficacy and efficiency of AI deployments. Subpar data quality elevates risks such as data breaches, security lapses, and other data-centric complications. These issues can inflate development expenses by necessitating the resolution of data inconsistencies, which in turn may lead to project delays and increased rework costs. Such disruptions can compromise the quality of AI solutions, diminishing customer satisfaction and eroding revenue and market share. Conversely, a commitment to high-quality data practices can assist SE management in curbing development costs, elevating product quality, enriching customer experience, and mitigating risks [18, 23].

Access to Data facilitates SE management by granting stakeholders insights into the data utilized in projects, development progression, and other pertinent details, aiding in identifying and mitigating risks associated with their chosen data for SOI. As businesses accumulate vast and diverse data sets, maintaining streamlined access becomes indispensable to prevent data landscapes from turning chaotic and complex [3]. Moreover, with tightening regulatory landscapes, such as the GDPR and the California Consumer Privacy Act (CCPA), adept data management, particularly regarding access, has gained paramount significance. Conversely, inefficient practices regarding data access can result in gaps in understanding data’s availability, quality, security measures, proprietorship, and overarching governance [18].

Privacy and Data are key elements in maintaining the integrity of AI systems, safeguarding against data breaches, and avoiding biased or discriminatory outcomes. AI systems often require access to data, including sensitive or personal information, that demands stringent protection measures. SE management stakeholders can play a vital role by incorporating strong privacy and data handling practices. These measures enable the ethical utilization of data, safeguarding against biased or prejudicial data sets and avoiding harm to individuals or groups. Wang et al. [38] point out that while data can provide invaluable benefits to organizations, it can also pose risks. High-profile cases like Meta (formerly Facebook) underscore the necessity for striking a balanced approach between exploiting data’s benefits and mitigating its associated risks, both from a social and regulatory standpoint.

System Security focuses on deploying security protocols like authentication and encryption to safeguard against unauthorized system or data access while ensuring that the system can quickly recover from any security breaches. The ultimate objective is to guarantee the system’s safe and reliable operation across diverse scenarios without harming users or society. Cheatham et al. [9] note that AI technology’s relative infancy means that SE management stakeholders often lack the refined understanding necessary to grasp societal, organizational, and individual risks fully. This lack of understanding can lead to underestimating potential dangers, overvaluing an organization’s ability to manage those risks, or mistakenly equating AI-specific risks with general software risks. To avoid or minimize unforeseen consequences, these stakeholders must enhance their expertise in AI-related risks and involve the entire organization in comprehending both the opportunities and responsibilities of AI technology.”

Fairness entails management practices of avoiding biased algorithms or data sets that may lead to discrimination or unfair treatment of certain groups [18]. It also means ensuring that AI systems design and development are supervised not to perpetuate or exacerbate societal inequalities. Berente et al. [5] explain that management stakeholders can ensure that the teams responsible for developing

and deploying AI systems are diverse regarding gender, race, and ethnicity to mitigate bias in decision-making. Diversity can help ensure that AI is designed and deployed fairly and ethically for all users, thereby increasing the adoption and acceptance of AI by a broader range of users.

4.2 Towards a Business Case for Ethical Requirements

To address RQ2 effectively, we introduce the Ethical Requirements Canvas, depicted in Fig. 2. This canvas serves to underline not just the importance but also the intrinsic value of ethical requirements, thereby constructing a business case for their integration. Business cases are essential for management to evaluate a project’s costs, benefits, risks, and alternatives, ensuring alignment with the organization’s strategic goals [40]. The Ethical Requirements Canvas serves as a practical instrument that not only integrates ethical considerations into management practices but also highlights their business value [28]. Consequently, the canvas provides a pragmatic method for aligning ethical requirements with the organization’s broader goals, articulating their significance and potential for adding value in business terms.

Ethics Requirements Canvas			Project Title	
Ethical Requirements	Key activities	Value proposition	Impact	Stakeholders
<i>Traceability</i>	<i>What activities are needed for implementation?</i>	<i>Promotion of Trustworthy AI practices</i>	<i>How does it impact your employees and customers?</i>	<i>Identify the categories of stakeholders that can be affected</i>
<i>Communication</i>				
<i>Data quality</i>	3	<i>Enhance stakeholder engagement and buy-in</i>	5	
<i>Access to data</i>	Key Resources	<i>Increased trust in AI technologies among users and society</i>	Societal Impact	
<i>Privacy and data</i>			<i>How does it impact society at large?</i>	
<i>System security</i>				
<i>Fairness</i>	4	<i>Increased sales</i>	6	2
1		9		
Cost			Benefits	
<i>What are potential negative impact that can result from a lack of implementation in our products or services such as reputational and financial impacts?</i>			<i>Value stream</i>	
<i>Discuss the negative impacts such as legal, privacy, environmental impacts, employment impacts etc.</i>			<i>Will customers need this and are they willing to pay?</i>	
<i>Redress costs from AI systems failing to operate or to be used as intended?</i>			<i>Which aspects of can be monetized?</i>	
<i>Adapted from The Business Model Canvas (Osterwalder, A., & Pigneur, Y. (2010))</i>			<i>Are customers willing to pay for it over what is obtainable</i>	
7			8	

Fig. 2. Ethical Requirements Canvas

Section one presents the ethical requirements identified through our research. It’s important to note that these requirements are displayed for reference and awareness, not for rigid adherence. Section two focuses on identifying the organization’s stakeholders. Here, SE management can discuss various categories of

stakeholders, such as human and non-human agents, different age groups, societal standing, and levels of vulnerability, among others. Section three outlines the essential business operations necessary to realize the value proposition of integrating ethical requirements. Section four lists the resources required for effective implementation. Sections five and six allow SE management to assess the societal, internal, and external impacts of incorporating these ethical parameters into their SOI. Section seven explores the financial, reputational, or otherwise costs associated with choosing to integrate or overlooking ethical requirements. Section eight evaluates the benefits and potential monetization of ethical requirements. Section nine illuminates the distinct advantages of ethical considerations, assisting in identifying vital initiatives that enhance the benefits of ethical requirements, potentially serving as critical determinants of success [7]. These benefits encompass elevating the organization to a Trustworthy AI business status, akin to the positive reputational impact observed in companies with sustainability initiatives. This can enhance stakeholder engagement-from the business being perceived as ethical and trustworthy-and potentially expanding market share and boosting profitability due to increased user trust. [7,27,28].

While the Ethical Requirements Canvas provides a systematic framework for visualizing and assessing ethical considerations, it may have inherent limitations. Its structured nature could risk simplifying complex ethical dilemmas, potentially fostering a compliance-centric mindset at the expense of cultivating a deeper ethical culture [31]. This approach risks satisfying only the minimum legal standards rather than aspiring to ethical excellence, which may lead to the marginalization of crucial ethical aspects [13,28,31]. Additionally, while adaptability is one of the canvas's strengths, it also poses challenges. Our research identified seven core ethical requirements, but their relevance and prioritization can differ significantly among organizations due to unique contextual factors, industry norms, and stakeholder expectations. Therefore, it is critical to balance adherence to industry standards with the strategic objectives of the organization when applying the canvas.

4.3 Limitation

A limitation inherent to our research is its specific focus on the marine transportation sector within Finland, potentially circumscribing the external validity and generalizability of our findings to other geographical contexts or industries experiencing AI-driven digital transformations. Despite this, we argue that our research lays a foundational framework that can be adapted and scrutinized in various settings [33].

For future studies, we plan to validate the Ethical Requirements Canvas via workshops with SE management teams and industry-wide surveys. These evaluations will not only gauge the canvas's usability and relevance but will also fine-tune its alignment with both organizational demands and ethical standards.

5 Conclusion

In this study, we have made three principal contributions. First, we compiled a comprehensive set of ethical requirements reflecting the perspectives of SE management stakeholders. Second, we presented a stakeholder-centric approach that is responsive to the challenges faced by the industry. Third, we introduced the “Ethical Requirements Canvas,” a novel tool designed to elucidate and integrate the value of ethical considerations into SE management practices. The canvas not only acts as an ethical roadmap for practitioners but can also facilitate risk management and promote judicious decision-making [28]. From an academic standpoint, our framework lays the groundwork for further inquiry into the integration of ethical requirements in AI and SE management, encouraging cross-disciplinary research and assessments of tool efficacy. On a practical level, our work supports SE managers in embedding ethical principles more deeply within their processes, thereby advocating for the development of trustworthy AI systems.

References

1. Agbese, M., Mohanani, R., Khan, A., Abrahamsson, P.: Implementing AI ethics: making sense of the ethical requirements, EASE 2023, pp. 62–71. Association for Computing Machinery, New York, NY, USA (2023)
2. Baker-Brunnbauer, J.: Tair framework canvas (2021)
3. Baker-Brunnbauer, J.: TAI Framework, pp. 97–127. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-18275-4_7
4. Barn, B.S.: Do you own a Volkswagen? Values as non-functional requirements. In: Bogdan, C., et al. (eds.) HESSD/HNSE -2016. LNCS, vol. 9856, pp. 151–162. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44902-9_10
5. Berente, N., Gu, B., Recker, J., Santhanam, R.: Managing artificial intelligence. *MIS Q.* **45**, 1433–1450 (2021)
6. Bobrowski, M., Marr, M., Yankelevich, D.: A software engineering view of data quality, February 1970
7. Boehm, B.: Value-based software engineering: reinventing. *ACM SIGSOFT Softw. Eng. Notes* **28**(2), 3 (2003)
8. Carter, M., Carter, C.: The creative business model canvas. *Soc. Enterp. J.* **16**(2), 141–158 (2020)
9. Cheatham, B., Javanmardian, K., Samandari, H.: Confronting the risks of artificial intelligence. *McKinsey Q.* **2**(38), 1–9 (2019)
10. Cohn, M.: *User Stories Applied: for Agile Software Development*. Addison-Wesley Professional, Boston (2004)
11. Consortium, D.: DSDM project framework. <https://www.agilebusiness.org/dsdm-project-framework/moscow-prioritisation.html>, January 2014
12. Dignum, V.: *Responsible Artificial Intelligence: how to Develop and use AI in a Responsible Way*. Springer, Cham (2019). <https://doi.org/10.1007/978-3-030-30371-6>
13. Elliott, K., et al.: Towards an equitable digital society: artificial intelligence (AI) and corporate digital responsibility (DDR). *Society* **58**(3), 179–188 (2021)

14. Freeman, P.: Software engineering body of knowledge (SWEBOK). In: International Conference on Software Engineering, vol. 23, pp. 693–696 (2001)
15. Guizzardi, R., Amaral, G., Guizzardi, G., Mylopoulos, J.: Ethical requirements for AI systems. In: Goutte, C., Zhu, X. (eds.) Canadian AI 2020. LNCS (LNAI), vol. 12109, pp. 251–256. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-47358-7_24
16. Halme, E., et al.: How to write ethical user stories? Impacts of the ECCOLA method. In: Gregory, P., Lassenius, C., Wang, X., Kruchten, P. (eds.) XP 2021. LNBP, vol. 419, pp. 36–52. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-78098-2_3
17. Hevner, A., March, S., Park, J., Ram, S.: Design science in information systems research. *MIS Q.* **28**(1), 75–105 (2004)
18. High-Level Expert Group on Artificial Intelligence (AI HLEG): Ethics guidelines for trustworthy AI. Technical Report, European Commission, April 2019
19. IEEE: Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems, version 2 (2017). https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf
20. IEEE: IEEE standard model process for addressing ethical concerns during system design. IEEE STD 7000–2021, pp. 1–82 (2021)
21. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* **1**(9), 389–399 (2019)
22. Lewis, D., Reijers, W., Pandit, H., Reijers, W.: Ethics canvas manual (2017)
23. Loshin, D.: The practitioner’s guide to data quality improvement. Elsevier (2010)
24. Mökander, J., Floridi, L.: Ethics-based auditing to develop trustworthy AI. *Mind. Mach.* **31**(2), 323–327 (2021)
25. Morley, J., Floridi, L., Kinsey, L., Elhalal, A.: From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Sci. Eng. Ethics* **26**(4), 2141–2168 (2020)
26. Müller, V.C.: Ethics of artificial intelligence and robotics (2020)
27. Open data institute: data ethics canvas (2019). <https://theodi.org/article/data-ethics-canvas/>
28. Osterwalder, A., Pigneur, Y.: Business Model Generation: a Handbook for Visionaries, Game Changers, and Challengers, vol. 1. Wiley, Hoboken (2010)
29. Papagiannidis, E., Enholm, I.M., Dremel, C., Mikalef, P., Krogstie, J.: Deploying AI governance practices: a revelatory case study. In: Dennehy, D., Griva, A., Pouloudi, N., Dwivedi, Y.K., Pappas, I., Mäntymäki, M. (eds.) I3E 2021. LNCS, vol. 12896, pp. 208–219. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-85447-8_19
30. Ray, P.P.: ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber-Phys. Syst.* **3**, 121–154 (2023)
31. Ricart, J., Casadesus-Masanell, R.: How to design a winning business model. *Harv. Bus. Rev.* **89**(1–2), 100–107 (2011)
32. Ryan, M., Stahl, B.C.: Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications. *J. Inf. Commun. Ethics Soc.* **19**(1), 61–86 (2020)
33. Shull, F., Singer, J., Sjøberg, D.I.: Guide to Advanced Empirical Software Engineering. Springer, London (2007). <https://doi.org/10.1007/978-1-84800-044-5>
34. Staron, M.: Action Research in Software Engineering. Springer, Cham (2020). <https://doi.org/10.1007/978-3-030-32610-4>

35. Swedberg, R.: Exploratory research. In: The production of knowledge: Enhancing Progress in Social Science, pp. 17–41 (2020)
36. 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)
37. Vidgen, R., Hindle, G., Randolph, I.: Exploring the ethical implications of business analytics with a business ethics canvas. *Eur. J. Oper. Res.* **281**(3), 491–501 (2020)
38. Wang, C., Zhang, N., Wang, C.: Managing privacy in the digital economy. *Fundam. Res.* **1**(5), 543–551 (2021)
39. Weber, R.P.: *Basic Content Analysis*, vol. 49. Sage, Thousand Oaks (1990)
40. Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A.: *Experimentation in Software Engineering*. Springer, Berlin (2012). <https://doi.org/10.1007/978-3-642-29044-2>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, 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 changes were made.

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.

