DIP-AI: A Discovery Framework for AI Innovation Projects

Mariana Crisostomo Martins
Instituto de Informática Federal
Universidade Federal de Goiás
Goiânia, GO, Brazil
maricrisostomo.martins@gmail.com

Lucas Elias Cardoso Rocha
Centro de Excelência em Inteligência Artificial
Goiânia, GO, Brazil
lucaseliascrocha@gmail.com

Lucas Cordeiro Romão
Departamento de Informática
Pontifícia Católica Universidade do Rio de Janeiro
Rio de Janeiro, RJ, Brazil
lc.romao98@gmail.com

Taciana Novo Kudo
Instituto de Informática Federal
Universidade Federal de Goiás
Goiânia, GO, Brazil
taciana@ufg.br

Marcos Kalinowski
Departamento de Informática
Pontifícia Católica Universidade do Rio de Janeiro
Rio de Janeiro, RJ, Brazil
kalinowski@inf.puc-rio.br

Renato de Freitas Bulcão-Neto
Instituto de Informática Federal
Universidade Federal de Goiás
Goiânia, GO, Brazil
rbulcao@ufg.br

ABSTRACT

Despite the increasing development of Artificial Intelligence (AI) systems, Requirements Engineering (RE) activities face challenges in this new data-intensive paradigm. We identified a lack of support for problem discovery within AI innovation projects. To address this, we propose and evaluate DIP-AI, a discovery framework tailored to guide early-stage exploration in such initiatives. Based on a literature review, our solution proposal combines elements of ISO 12207, 5338, and Design Thinking to support the discovery of AI innovation projects, aiming at promoting higher quality deliveries and stakeholder satisfaction. We evaluated DIP-AI in an industry-academia collaboration (IAC) case study of an AI innovation project, in which participants applied DIP-AI to the discovery phase in practice and provided their perceptions about the approach's problem discovery capability, acceptance, and suggestions. The results indicate that DIP-AI is relevant and useful, particularly in facilitating problem discovery in AI projects. This research contributes to academia by sharing DIP-AI as a framework for AI problem discovery. For industry, we discuss the use of this framework in a real IAC program that develops AI innovation projects.

KEYWORDS

Requirements Engineering, Discovery, AI Systems, Innovation Projects.

1 Introduction

The increasing adoption of Artificial Intelligence (AI) components in software raises concerns regarding the quality of these components. Many factors have contributed to the widespread integration of Machine Learning (ML) components into traditional software. Examples of such systems can be found in healthcare [27], finance [17], education [28], and others.

Difficulty in communication and high customer expectations are aspects commonly addressed in Requirements Engineering (RE) and identified as particularly challenging in ML projects [25]. These systems often involve ambiguous requirements, emergent quality attributes such as explainability, and inherently iterative and experimental development cycles. Despite its critical role in quality assurance, current RE practices offer limited guidance for ML-based systems [1]

Martins et al. [30] conducted a study on the state of RE practice in ML projects within an innovation-driven environment, focusing on research, development, and innovation (RDI) projects. Through interviews with project coordinators, they identified several challenges, including problem comprehension, deriving insights from data, maintaining effective communication, and aligning technical and domain knowledge to manage feasibility constraints challenges. These findings are aligned with the findings from the survey by Alves *et al.* [5], which also identified difficulties in problem understanding, managing customer expectations, and aligning requirements with data as key challenges for the improvement of the quality of these systems.

The development with AI components represents a paradigm change compared to traditional software development. AI systems present AI models that have data-based behavior while in traditional software the behavior is specified according to user needs and business rules. RE is a challenge in this scenario, being critical and directly affecting the quality of these systems [16, 40]. The traditional practices of RE are not well aligned with AI development [16, 24, 46].

Marban *et al.* [29] present a comparison between CRISP-DM [11] and the methodologies proposed by Fayyad *et al.* [14], Cabena *et al.* [9], and Anand and Bucher [6]. These methodologies have in common a *Discovery* phase, comprising the identification of human resources, understanding the business, determining business objectives, and developing and understanding the application domain. Although they are aligned with the objectives of ISO 12207, there is no clear mapping to the aspects of the ISO requirements. The identification and characterization of the business becomes a challenge to be addressed by the project team, usually by the data scientist. In practice, this stakeholder typically conducts RE-related activities [5, 33]. However, these activities would be better developed by a specialist with the support of data scientists and other stakeholders [42]. In fact, to identify a business and characterize it in search of points that reinforce the need for AI solutions for the business, it is necessary to have a direction to support this activity, together with other stakeholders.

This paper proposes a framework for discovery in AI innovation projects that helps to understand the business, its objectives, and

its problems, as well as to prepare for the project definition. This framework guides requirements analysts in understanding the organization, problems, data, and identifying characteristics and their distribution, which is essential to achieve a high-quality AI model [45]. This solution facilitates understanding of the context and domain [23], enhances stakeholder collaboration [25], and enables anticipation and structured discussions on model degradation [44], perception regarding quality, non-functional requirements (NFR), and trade-offs [1, 19, 41]. We developed the proposed framework based on the Technology Transfer Model (TTM) [36], incorporating foundational elements from ISO 12207, 5338, and Design Thinking.

To assess the effectiveness of the proposed framework and gather insights for its improvement, we conducted a case study following the guidelines outlined by Runeson and Höst [38]. The study was carried out within the context of a research and development (R&D) initiative, independent of the authors' institution, and involved a real industrial partner engaged in the development of a minimum viable product (MVP). The evaluation of the DIP-AI framework was structured as an evaluative case study, focusing on the perspectives of participants involved in the problem discovery phase.

We designed a structured questionnaire aimed to capture participants' perceptions regarding the problem discovery capability, usability, and utility of the framework, with the quantitative items grounded in the constructs of the Technology Acceptance Model (TAM) [39]. We also collected suggestions with additional open-text answer options.

As part of the case study, DIP-AI was successfully applied in practice, helping to understand the business objectives and problems, providing the basis for subsequently defining an MVP of an AI product that was successfully developed and deployed at the industrial partner. The main findings of the quantitative analysis indicate that the proposed approach is useful, relevant, and effective in supporting the discovery of AI innovation projects. Qualitative analysis of participants' feedback revealed additional suggestions for enhancement, particularly regarding the need for additional support for applying the framework.

This paper is organized as follows. Section 2 describes the related work. Section **??** presents the background of the proposed framework. Section 4 presents the proposed framework and describes the method for defining the framework. Section 5 describes the study case design. Section **??** provides the study case results. Section **??** presents the discussion of the results. Section 8 discusses threats to validity. Finally, Section 9 concludes this work, highlighting its contributions and future work.

2 Related Work

This section reviews related approaches to RE for AI systems. While existing works define requirements and perspectives (e.g., PerSpecML [42]) or combine methodologies like CRISP-DM, DT, and Lean [2], none focus specifically on a structured Discovery phase for AI projects.

PerSpecML organizes RE-related concerns into five perspectives (goals, user experience, infrastructure, model, and data) and supports their specification through guided questions. Our approach complements it by preceding the specification with structured problem discovery and feasibility analysis. Ahmed *et al.* propose an integrated methodology (business, data, product), but do not address specific aspects of development defined in the RE process and relevant for quality development.

The AI Canvas [12], the ML Canvas [13], and Deep Learning (DL) Canvas [35] are canvases that present fields with context application, key evaluation metrics, features, model development, and data aspects (distribution, source, and logistics). All of these fields are very general and, therefore, depend on the expertise of the filler to obtain the necessary information. In addition, aspects regarding the decision of which problem(s) will be solved, data availability, model consumption method, etc. are not addressed in the canvas and can strongly impact the project.

Our framework addresses these gaps by supporting early-stage discovery-related activities, such as problem identification, KPI definition, feasibility assessment, and stakeholder alignment, to enable more informed and quality-driven AI innovation projects. Hence, we emphasize that the unique aspect of our contribution is a proposal that brings together aspects of an AI project that can be considered from the project's discovery stage and will assist in decision-making so that the project has higher quality and meets stakeholders' expectations.

3 Background

3.1 ISO/IEC 12207:2017 and ISO/IEC 5338:2023

ISO/IEC/IEEE 12207:2017 [26], entitled "Systems and Software Engineering - Software Life Cycle Processes", is an international standard that defines a process for software engineering (SE) throughout the life cycle of a software. It is jointly published by the International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC), and the Institute of Electrical and Electronics Engineers (IEEE). The standard is widely used across industries to ensure that complex systems are efficiently and effectively designed, developed, operated, and maintained.

The standard covers all phases of the software life cycle and defines a set of processes and activities necessary for SE. These processes include i) agreement; ii) project organization; iii) technical design; and iv) technical support. It is designed to be compatible with and complementary to other standards, such as ISO/IEC/IEEE 15288:2023, which focuses on systems engineering, and ISO/IEC 5338:2023 [25], which introduces AI-specific processes and updates requirements engineering practices for AI systems.

Next, we present key points about ISO/IEC/IEEE 12207:2017, essential for the context of this proposal:

- *Requirements Engineering*: emphasizes the importance of identifying and documenting software requirements to guide software design and development, with dedicated processes for requirement elicitation and specification.
- *Risk Management*: addresses risk identification and mitigation, along with quality assurance, throughout the development process to ensure software reliability.
- *Adaptability*: offers flexibility to suit different types of software systems and organizational contexts, making it suitable for innovative and AI-driven projects.
- *Stakeholder involvement*: highlights the need for continuous engagement with stakeholders to ensure their expectations are met throughout the life cycle.

- *Specification and Traceability*: stresses the importance of thorough documentation and regular reviews to maintain alignment with requirements and ensure expected functionality.

In this work, we adopt this standard to support the definition of the requirements engineering process and associated artifacts, aiming to ensure problem identification, quality, and traceability in the development of AI-based systems.

3.2 Design Thinking

Some works highlight the use of Design Thinking (DT) and RE [3, 20–22, 34]. It is a structured approach to problem-solving and has been used to develop innovative products that explores needs and integrates an agile and flexible environment to solve complex and ill-defined problems.

After examining the two approaches, we identified that ISO/IEC 12207 and DT can complement each other in proposing support for the discovery of AI innovation projects. DT follows a structured process represented by the Double Diamond model, which alternates between divergent and convergent thinking phases—first to explore and define the problem, then to ideate and prototype solutions.[18]. The focus of our proposal is on the first diamond to support problem identification. Identifying problems and detailing them is challenging, as is establishing communication with stakeholders. Other contributions already support RE ideation [4] and specification activities [42].

4 Methodological Path

Our proposal was conceived and refined based on the Technology Transfer Model (TTM) [36], which supports the development of artifacts (e.g., models, methods, and processes) to address specific challenges.

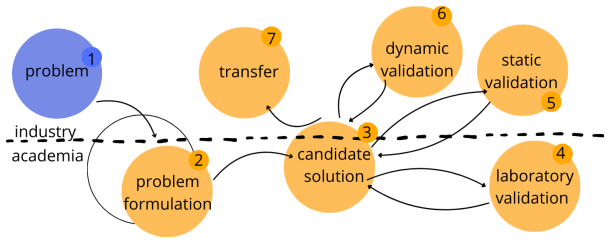

Figure 1: Transfer Technology Model adapted by [36].

As illustrated in Figure 1, TTM begins with the identification of a real-world industry need, supported by findings in the literature. Based on this need, a candidate solution is designed and iteratively evaluated in various scenarios. Each evaluation cycle informs refinements to the artifact, enhancing its relevance and value for both research and practice. The next sections detail how the TTM was applied to conceive DIP-AI.

4.1 Problem identification

We conducted interviews with coordinators of several of AI and innovation projects. We discovered the difficulty in understanding the underlying problem associated with data-driven tasks, ensuring effective communication among stakeholders, and aligning domain-specific and technical knowledge[30]. These challenges were also found by Alves et al. [5].

4.2 Problem formulation

We conducted a tertiary study, that is, a review of secondary studies [32]. We searched for studies that related RE and AI in databases such as ACM DL, IEEE Xplore, Scopus, Engineering Village, Wiley, and Web of Science. We obtained 1280 studies and included 9 for full reading and data extraction. We identified gaps in RE for AI systems, particularly in the early stages of problem and data identification, communication and collaboration with stakeholders, metrics and acceptance criteria identification, NFR and trade-offs analysis, and traceability.

Given this context, we formulated the following research question: **"How can we support problem discovery to enhance the effectiveness of AI innovation projects?"**. In response, we proposed a candidate solution described in the following section.

4.3 Candidate Solution: DIP-AI

4.3.1 DIP-AI Construction Process. We studied the RE-related ISOs technical processes. We listed the activities and tasks of each process and highlighted what we considered essential for executing the RE phase so that they would not be suppressed. We also defined activities with experienced requirements engineers and the support of a data scientist. The activities were discussed iteratively over four weeks.

Regarding the *business analysis* process, we included the following activities: (i) define the scope of the problem and analyze complaints, (ii) characterize the solution space and identify alternative solutions, (iii) evaluate alternative solutions, and (iv) maintain traceability. For the *stakeholders' needs definition* process, we defined the following activities: (i) identify stakeholders; (ii) define needs, context of use, and scenarios; (iii) classify and prioritize needs; (iv) identify restrictions and relationships with non-functional requirements; (v) analyze the set of requirements and define performance measures; (vi) discuss and give feedback; and (vii) obtain agreement and maintain traceability.

Concerning the *system requirements definition* process, we chose the activities as follows: (i) define the functional boundary; (ii) identify modes of operation, implementation restrictions, and risks, and define requirements; (iii) analyze requirements; and (iv) obtain agreement and maintain traceability. Finally, we defined the following activities for the *system analysis* process: (i) identify contexts and assumptions, analyze the results, establish conclusions and recommendations, record results, and (ii) maintain traceability. We believe that a simplified guide process can aid experimental and iterative AI and innovation development.

After defining the essential activities, we aligned them with the DT Double Diamond model. In the empathize phase, we positioned activities from the business analysis process. The define phase encompasses tasks related to the stakeholder needs definition process. The ideate phase integrates activities from the system requirements definition process. The system analysis process spans across all

three phases, supporting iterative refinement and stakeholder alignment. Finally, the prototype phase involves the development, evaluation, and refinement of ML-enabled prototypes until a solution that meets stakeholder expectations is achieved [31].

Each DT phase also has expected outcomes aligned with ISO standards. The empathize phase aims to uncover user and business needs. The definition phase results in the articulation of challenges, problems, and evaluation criteria.

For each phase, we specify relevant information and subprocesses suitable for AI innovation projects. Based on this process, we formulated guiding questions to assist in problem identification, solution analysis, and data-driven stakeholder engagement critical to this development context. Table 1 presents the relationship between the main constructs of the canvas and ISO 12207 processes.

The 5W2H management tool has been used to make the planning and execution of projects, tasks, and actions clearer, ensuring that all relevant information is provided, assigned, and understood by all involved [15]. We identified that to understand the problem we needed to answer some basic questions such as: **What** are the businesses involved? **How** are they carried out? **Who** are the stakeholders? **Where** do the processes take place in the organization? **How** do these processes happen and **why**? **What** are the costs and impacts of these activities? Aligned with the 5W2H questions. These questions help better characterize the organization's existing processes and problems. This allows for greater insight into the project management landscape and better alignment with stakeholders.

The GUT Matrix prioritizes problems based on Gravity, Urgency, and Trend, helping guide decision-making in dynamic project contexts [15]. This GUT matrix classifies problems to establish ordered priorities to deal with problems efficiently. Therefore, the processes, problems, and tasks are prioritized (using the GUT matrix). Then the feasibility analysis is carried out by collecting information such as input and output data, algorithms, metrics, references, etc.

We also identified the importance of understanding the task data associated with the problem as part of feasibility, as this determines whether an AI solution is applicable to the project. We refined DIP-AI by discussing it with project coordinators. The current version of the DIP-AI framework canvas is presented in Figure 2 and details about its structure follow.

4.3.2 DIP-AI Framework Canvas. The proposed canvas is structured into three parts, following the first diamond of the Design Thinking model. The first part corresponds to the empathy and immersion phase, aimed at understanding the client's context and pain points. The second part focuses on defining and prioritizing problems and project objectives. The third part refers to ideation, where possible solutions for the selected priority problem are outlined, followed by a feasibility analysis.

Stakeholder involvement—an acknowledged challenge in AI projects [31]—is central to using the canvas properly. In the first phase, fields are completed primarily by the Business Owner (BO), who provides context, organizational background, existing challenges, and identifies other relevant stakeholders. In the second phase, the Domain Expert (DE) joins the BO to characterize and prioritize organizational processes, problems, gaps, and opportunities. At the end of this phase, the Data Scientist (DS) identifies which of the prioritized elements correspond to data-related tasks. In the third phase, the DS answers more advanced questions to evaluate the feasibility of developing AI-based solutions for the selected tasks.

By guiding this multi-stakeholder collaboration and addressing specific AI-related challenges, the canvas offers a practical approach tailored for AI innovation projects [5, 30]. It was designed to promote an iterative and adaptable approach suited for AI innovation projects, aimed at facilitating communication through its visual structure, enabling the alignment of business objectives and technical feasibility with data, and supporting the understanding and prioritization of problems and processes.

4.4 Validation

The objective of these stages is to progressively refine the candidate solution to ensure its suitability for industry adoption. Following the Technology Transfer Model (TTM), **we carried out three evaluations in distinct contexts:** (i) an academic validation with graduate students in Software Engineering for AI, who used the approach to identify problems in their final course projects; (ii) a static validation in an industrial setting, where RDI professionals retrospectively applied the approach to real project artifacts—this type of validation is conducted offline, with real practitioners and data, but outside the context of active development; and (iii) a dynamic validation, embedded in a real-time industrial case study within an ongoing project. These iterative assessments were designed to promote early issue identification, enhance user satisfaction, foster continuous improvement, and build trust in the solution's applicability. An approved ethics review was conducted prior to performing validations involving human participants. Subsequently, a pilot study was carried out with five graduate students to assess the validation procedure and data analysis methods, as well as to identify potential issues [1].

4.5 Release the solution

In the future, we plan to implement the proposed improvements and develop documentation to effectively transfer the technology.

5 Study Case Design

This investigation is structured as an evaluative case study involving participants actively engaged in a discovery phase to identify and characterize a viable problem for the project. The methodological approach was guided by the case study research principles outlined by Runeson *et al.* [38].

5.1 Goal and Research Questions

We followed the Goal-Question-Metric (GQM) goal definition template [8] to define our research goal as follows: ***Analyze* DIP-AI with the *purpose* of characterization *with respect to* the problem discovery capability, overall acceptance, and criticisms and suggestions for improvement, from the point of view of participants in the context of IAC AI innovation projets.**

Based on this objective, we formulated the following Research Questions (RQs):

[1] CAAE: 81475724.4.0000.5083

Table 1: Relation between ISO/IEC/IEEE 12207 modified by ISO/IEC/IEEE 5338 processes and DIP-AI Canvas

Process/Fields Goups	Question about context projects	5W2H	GUT Matrix	ISO 31000	Fesibility Analysis
Business or mission analysis process	Define scope of the problem and characterize the solution space				Evaluate classes of alternative solutions
Stakeholder needs and requirements definition process		Define context of use	Prioritize and classify needs	Identify constraints	
System requirement definition process		Identify and define needs and justifications	Prioritize needs	Identify risk-related requirements	
System analysis process					Define strategy and perform analysis

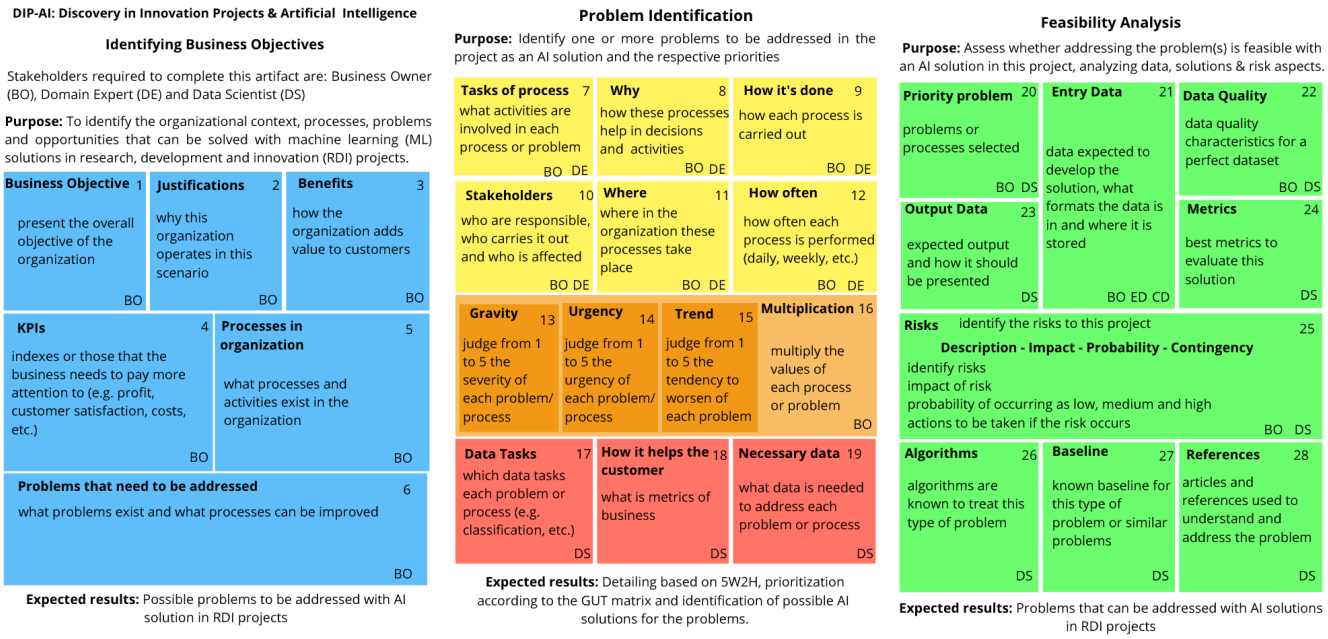

Figure 2: DIP-AI Canvas.

- **RQ1: How well does DIP-AI contribute to problem discovery in AI innovation projects?** To answer this question, we closely followed the experience of applying DIP-AI, analyzed the completed DIP-AI canvas artifact, specific follow-up questions, and considered the concrete result of the project (which, at the time of writing, has been completed and deployed at the industrial partner).
- **RQ2: How well is DIP-AI accepted in practice?** To answer this question, the follow-up questionnaire was designed based on the Technology Acceptance Model (TAM 3) [39].
- **RQ3: How can DIP-AI be further improved?** To answer this question, we qualitatively analyzed open-text responses to the follow-up questionnaire regarding criticisms and improvement suggestions.

5.2 Validation Context

The context of this study is an industry-academia collaboration innovation project initiative designed to enhance the education of undergraduate and postgraduate students. In this initiative, participants play a central role in the design and development of innovative AI-based technology solutions aimed at addressing concrete business challenges of real industry partners [37]. The industrial partner involved in the particular case study is a technology company that develops digital solutions for multiple clients, with a product portfolio focused on technological security across various industry sectors.

Problem discovery is an important part of an AI innovation project, enabling the identification of problems rooted in the genuine needs of the industrial partner. The project in question was suitable for our case study as it did not yet have a clearly defined problem. We spent a month on-site accompanying the participants whilecarrying out the Discovery phase of the project with the support of DIP-AI. We participated in the onboarding meeting with the customer, daily meetings, and applied the questionnaire.

Regarding the profile of the participants, the project team had eight undergraduate students and two doctoral students. All participants had at least one year of prior experience in AI innovation projects, having been involved in between two and six such projects. Six of the participants declared themselves experienced and skilled in identifying problems in AI projects. More details about the participants' profiles are shared with other materials from this evaluation [2]. Figure 3 illustrates the participants in the context of the case study and the DIP-AI framework used in the background.

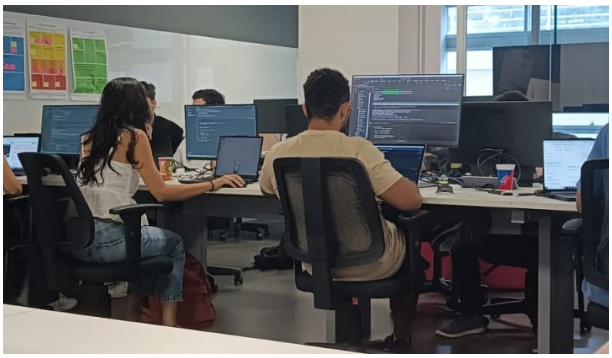

Figure 3: Participants in the Case Study Context and the Use of DIP-AI in the background.

Figure 4 illustrates the DIP-AI framework canvas completed during the case study. The canvas was first introduced to the project coordinators on January 15, 2025, when the constructs employed in the canvas, the potential benefits of its evaluation, and the components of the artifact were presented. On January 22, 2025, the first author conducted a presentation to the participants, providing an overview of the canvas and explaining each field in detail. During this session, participants posed questions about its use and engaged in discussions about the relevance of the canvas and how it might help to mitigate challenges in their project. Each presentation lasted approximately one hour.

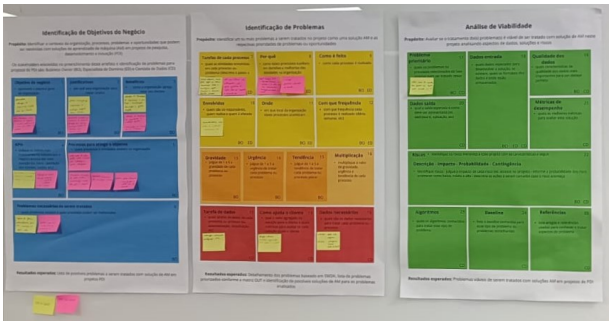

Figure 4: DIP-AI completed by participants in the case study.

In the subsequent days, the first author documented difficulties reported by participants in their respective project that could potentially be addressed through the proposed approach. Additional contextual information about the industrial partner and the project challenge was also gathered. On January 29, 2025, the project objective was revised in collaboration with the partner company and the innovation unit due to feasibility concerns. As a result, the canvas was completed for a second time on January 30, 2025, between 3:30 p.m. and 6:00 p.m. During this session, participants identified and prioritized multiple project-related tasks using the GUT Matrix.

In the days that followed, participants continued to raise questions about specific canvas fields, although they did not resume completion immediately. The process was resumed on February 6, 2025, from 3:20 p.m. to 4:45 p.m., at which point participants were better prepared to engage with the artifact. Having met with company representatives to clarify uncertainties and deepened their understanding of the problem, they completed all remaining fields.

5.3 Data Collection

Data collected throughout the experience of the participants applying DIP-AI involved the completed DIP-AI canvas artifact, the follow-up questionnaire, and information on the concrete result of the project.

For RQ1 (contributions to discovery) we considered the overall problem discovery experience, questions about perceptions of contributions of the artifact (questions that we created and that are shown in group 1 in Figure 5), and the open question "In the projects in which you participated, what were the difficulties in identifying the problem or machine learning task to be addressed in that project? Could they be mitigated by using this artifact?".

For RQ2 (acceptance), we considered the answers to the TAM-based [39] questions. Considering the research objectives and context, the study focused on Technology Acceptance Model (TAM) constructs related to perceived usefulness—specifically job relevance, output quality, and result demonstrability—and perceived ease of use, including computer self-efficacy, perception of external control, perceived enjoyment, and objective usability. To ensure methodological rigor, the original semantics of the selected constructs were preserved, and internal reliability was assessed using Cronbach's Alpha, yielding a value of 0.95, which indicates excellent reliability. We selected eight of the TAM 3 constructs (questions of groups 2 to 9 in Figure 5), which are: usefulness, ease of use, ability to perform the task using the artifact, control of the artifact to perform the task, pleasure in using the artifact, perceived relevance of the artifact, output quality and demonstrability, and intention to use the artifact.

Finally, for RQ3, we considered the answers to the following open questions:

- What limitations, recommendations, or opportunities for improvement do you identify in this artifact?
- Do you consider that any questions or fields are missing from the artifact?
- Do you find the current structure and organization of the artifact to be clear and appropriate?

5.4 Analysis Procedures

For the quantitative analysis of the TAM questions, we performed a reliability analysis using Cronbach's Alpha. Reliability indicates how reliably or accurately a questionnaire or test measures a true

[2] https://zenodo.org/records/15732851

value. Cronbach's alpha is given by $\alpha = \frac{N\bar{c}}{\bar{v}+(N-1)\bar{c}}$, where N represents the number of items, $\bar{c}$ is the average inter-item covariance, and $\bar{v}$ is the average variance. The quantitative analysis was conducted by aggregating the number of responses for each questionnaire item, based on a seven-point Likert scale to assess the perceived contribution of the framework.

The qualitative analysis of open questions was conducted using the Atlas.TI tool [7], applying open and axial coding procedures from Grounded Theory [10].The method we used was to gather participants' responses to each open-ended question in the questionnaire. We then analyzed the contributions, coded each response, grouped the responses, and established possible relationships. Since we were unable to reach data saturation due to the small sample of participants, we did not perform selective coding.

6 Results

In Figure 5, we present a heat map that represents the responses of the 10 participants to the 36 Likert scale items (the five created ones for artifact contribution, and the 31 from the 8 selected TAM 3 constructs) that form the perception questionnaire of this research. The questions are numbered in the form X.Y where X represents the group of questions and Y enumerates the questions in a group. The values in each field of the heat map represent the number of responses to a given question. We obtained a Cronbach's Alpha of 0.950, which means that the internal consistency of the questionnaire can be considered excellent. Hereafter, we answer each of the research questions.

6.1 RQ1: DIP-AI's contribution to problem discovery

Regarding RQ1, considering the overall discovery experience, the team filled the DIP-AI canvas and used it to guide the identification of the problem. In fact, it supported a strategic redefinition of the problem by analyzing risks and opportunities early. The team started working on understanding a particular problem, but then identified risks and a more promising alternative, changing direction. It is noteworthy that at the time of writing, an MVP solving the identified problem was successfully developed and deployed at the industry partner.

Furthermore, based on the first group of questions, it is possible to observe that participants reported that use of the DIP-AI canvas positively influenced project execution by optimizing resources and supporting the achievement of objectives, particularly because problem identification was conducted accurately, efficiently, and effectively.

Additionally, participants acknowledged that the use of the DIP-AI artifact helped address common challenges encountered during the problem discovery phase in innovation and AI-related projects. However, two participants disagreed about the artifact's usefulness for individuals identifying a problem in a project for the first time.

Participants' answers to the question "In the projects in which you participated, what were the difficulties in identifying the problem or machine learning task to be addressed in that project? Could they be mitigated by using this artifact?".

- P3: "I believe that using this artifact right from the beginning could help us look at data and request possible pending issues that may exist."
- P6: "I believe that using this artifact the difficulties could be mitigated"
- P8: "Lack of clarity of processes related to the problem and there is a need to structure the necessary data in a clear way."
- P7: "Lack of clarity in communication with the customer and in the company's processes are difficulties that could be mitigated, since the artifact would help the team prioritize elements of an AI task."

RQ1 Answer: The use of DIP-AI was suitable, and it was considered to contribute positively to the discovery phase of the AI innovation project.

6.2 RQ2: DIP-AI overall acceptance

Participants expressed positive perceptions regarding the **usefulness** of the artifact (as reflected in the responses from Group 2). They considered the artifact to be beneficial in several ways: it enhances productivity, increases assertiveness in decision-making, and supports effective problem identification.

Regarding **ease of use** (Group 3), four of the participants reported that the artifact requires a lot of mental effort. And only one student reported that he did not find it easy to use the artifact for what he wanted.

Regarding the group of questions (Group 4) related to the **ability to identify problems** using the artifact, two participants reported that they would have difficulty completing the activity without guidance. Additionally, one participant indicated that relying solely on a support resource or a brief explanation would not be sufficient to carry out the task effectively. These perceptions align with comments from other participants who noted that the artifact is not entirely intuitive or easy to use.

Regarding the perception of **controlling the artifact** (Group 5) to perform the activity, three participants disagreed with having control over the artifact. Perhaps because they consider it not a very customizable tool.

Regarding the dimension of **pleasure** (Group 6) in performing the activity with the artifact, one participant reported not experiencing any sense of enjoyment, and three others stated that they did not find the task engaging or fun. These responses are consistent with the broader perception that the artifact is not particularly easy to use, as it is associated with a cognitively demanding activity that requires significant mental effort.

No negative perceptions were reported regarding the **relevance** of the artifact (Group 7) or the importance of its use in supporting tasks related to AI project development. Similarly, there were no negative comments concerning the demonstrability and quality of the artifact's results (Group 8). On the contrary, participants emphasized that the artifact contributes positively to stakeholder communication, reinforcing its practical value in project contexts.

Finally, nine responses were positive regarding the **intention to use** (Group 9) the artifact in the future.

RQ2 Answer: The use of DIP-AI was well accepted by agile team involved in innovation project.

Figure 5: Heat Map of Participant's Perception (TAM 3-based)

6.3 RQ3: DIP-AI criticisms and suggestions

Participants generally reported that the artifact was well-structured and adequately organized, with no missing fields. They also mentioned that the use of the framework could help mitigate challenges previously encountered in other projects.

Participants indicated that DIP-AI is particularly valuable during the discovery phase of AI innovation projects. Three responses highlighted that the framework contributed to understanding the problem and facilitated communication with stakeholders. Additionally, four participants noted that the DIP-AI canvas played a key role in stakeholder discussions concerning the availability of data, an essential aspect for the success of such projects. Participants also identified specific advantages of using the artifacts: one response emphasized the artifact's role in task prioritization, and three others mentioned that the canvas helped reduce common difficulties in innovation and AI initiatives.

Despite the positive feedback, participants also offered several suggestions for improvement. One participant noted that the artifact is extensive and proposed completing it in iterative rounds. Another suggested that some fields could be merged to reduce the overall complexity. Additional suggestions included the inclusion of a glossary to clarify terms and acronyms, as well as providing guidance to assist stakeholders in completing the artifact. One participant recommended enriching the "priority problem" field by linking it to problem characterization — specifically, incorporating information from the GUT matrix and task data into the analysis of findings. Another participant proposed integrating the canvas with agile methodologies, such as Scrum, to periodically reflect on the project's objectives during each sprint. Finally, two respondents suggested that the canvas would benefit from tool support to enhance user interaction and usability.

- P3: "I believe that using this artifact right from the beginning could help us look at data and request possible pending issues that may exist."
- P6: "I would add, along with the algorithms, the characterization of the priority problem. For example: it is a classification problem, regression, supervised learning, unsupervised, it will involve Deep Learning, LLMs."
- P8: "I would like some concepts of the artifact to be explained through a glossary, because if we did not have help, it would be confusing to understand some terms."
- P7: "I would use the artifact as an aid to SCRUM, just like KANBAN. Filling out the artifact in "rounds" is one of its greatest values: giving a temporal notion of the team's perception of the project as a whole. It may even have too many fields, which could lead to exhaustion for those using the artifact and cut short the creative innovation process a bit by keeping the user too tied to a (possibly) long and closed step-by-step process. In my opinion, the artifact could benefit a lot from tool support with an inviting and clean UX."
- P10: "During the dynamic, I felt that some fields were confused with each other, and perhaps, without the help of the researcher, I would not have properly understood what differentiated them. For example, the GUT matrix was a little confusing to me at first, although it was very useful. I found it more practical to use it virtually."

RQ3 Answer: DIP-AI contributed to problem discovery, communicating with stakeholders, and prioritization of problems. We also received suggestions to make the canvas more user-friendly, such as adding a glossary and tool support.

7 Discussion

DIP-AI is a proposal aligned with the RE phase applied to AI projects. In AI project development processes, there is concern with the Discovery phase, where the problem is identified and characterized; the data is discovered, and the project is prepared for the other phases. Project coordinators pointed out the difficulty in identifying the problem and in building and maintaining this documentation for their projects [30]. This difficulty was also pointed in studies on the state of the art and practice [5, 43]. The approach was developed to help discovering and selecting problems to be addressed in a project. DIP-AI was useful in identifying the problem to be addressed in the project. The AI product developed and delivered within the case study was based on a WhatsApp user interface and employed LLMs for fluent communication and an ML classification model to identify potentially unsafe situations, also proactively suggesting protective actions to users. The engine is capable of interpreting multiple input formats provided through WhatsApp (QR codes, images, locations, and text messages).

In the evaluative case study, the proposed framework was perceived as both useful and relevant. The quantitative analysis supports this perception, indicating that the artifact effectively aids in problem identification. These findings are corroborated by the qualitative analysis, which reinforces the positive evaluations observed in the quantitative data. The results suggest that the framework facilitates communication and assists in collecting project-related information that may signal potential future risks, such as the viability and availability of data and models. However, perceptions regarding the ease of use were not uniformly positive. The qualitative feedback provided valuable insights on how to enhance the proposal's usability. Suggestions include simplifying the completion process through the use of a glossary, guided assistance, or iterative completion rounds.

8 Threats to validity

In this section, we discuss the threats to the validity of our study, according to the categories presented by Runeson *et al.* [38].

In terms of **construct validity**, potential challenges include the influence of assessment apprehension, which may mobilize participants to avoid negative feedback. To mitigate this threat, the questionnaire was administered anonymously. In addition, hypothesis guessing emerges as a potential threat, where participants may have provided overly positive responses, anticipating the research being conducted. To minimize this problem, we clarified that the questionnaire would be used to investigate ways to understand and improve the experience of identifying problems in projects. Participants were encouraged to be honest and critical in order to improve the approach in the future.

For **internal validity**, issues such as diffusion or imitation of treatments come to the fore. Some participants may have sought information about the motivation for some responses from their peers, compromising the integrity of our study. To minimize this

threat, we included open-ended questions at the end of the questionnaire to complement the quantitative analysis of the results and better understand participants' perceptions when answering each question. Furthermore, we applied the questionnaire individually and synchronously to avoid the influence of comments among participants.

Regarding **external validity**, we faced the selection and treatment interaction due to the non-representative sample achieved by our study. The participants chosen for the research were currently active in AI innovation projects and had already performed problem identification in previous projects. Despite this, we recognize the importance of carrying out more case study evaluations with different participant profiles and projects for generalization.

Finally, regarding **reliability**, we conducted a single case study. However, as expected for case studies, we carefully described the context, and our case study fills a gap in the literature by identifying problems of innovation and AI projects in an IAC context with a large Brazilian security company. To improve the reliability of our analyses, we peer-reviewed all of our analysis procedures involving two independent researchers and made all data available online for audit in our open science repository.

9 Final Remarks

This paper presents DIP-AI, an approach developed to identify and prioritize key problems to be addressed in artificial intelligence innovation projects. The approach empowers requirements engineers — working collaboratively with business owners, domain experts, and data scientists — to uncover opportunities for applying AI to solve business problems. We introduced the DIP-AI canvas with 28 fields for identifying the business objectives, characterizing problems and opportunities, and analyzing feasibility.

We carried out a case study in the discovery phase of an ongoing project with ten team members who had previous experience with AI innovation projects. We held workshops for the participants to fill out the DIP-AI canvas, and the participants had the opportunity to ask questions about filling out the canvas and collecting information from customers.

The DIP-AI canvas was reviewed by data scientists and innovation project coordinators who provided feedback. The perceptions collected from the questionnaires were positive regarding the usefulness and relevance of the canvas. Despite this, we received negative feedback regarding the ease of use and pleasure in using the artifact. These perceptions were corroborated in the qualitative analysis, where the participants indicated that the proposal could benefit from a glossary and tool support for filling out the canvas.

The case study demonstrates that the DIP-AI canvas supports the formulation of innovative artificial intelligence solutions by aligning organizational objectives with process mapping and customer problem identification. It enhances stakeholder communication, guides the exploration process from pain point identification to feasibility assessment, and offers insights into practitioner needs, thereby informing future research aligned with the requirements engineering phases defined in ISO/IEC 12207 and ISO/IEC 5338. We also offer practical implications for industry by presenting a successful application of the framework for problem identification in a real AI innovation project developed for a large technology company.

As future work, we plan to conduct additional case studies encompassing a broader and more diverse range of artificial intelligence innovation projects. Furthermore, we intend to design and evaluate a software tool to facilitate the application of DIP-AI and to assess its usability and contributions through focus group sessions with domain experts.

ACKNOWLEDGMENTS

We thank ExACTa, including the program sponsors, business managers, and especially the Digital Innovation team, for their support in establishing a fruitful collaboration between industry and academia. We also thank the students, mentors, and faculty enrolled in the program for their commitment to making this collaborative educational program a valuable learning experience for all involved. We thank the Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES) for its financial support—Financial Code 001.

REFERENCES

[1] Khlood Ahmad, Mohamed Abdelrazek, Chetan Arora, Muneera Bano, and John Grundy. 2022. Requirements Engineering for Artificial Intelligence Systems: A Systematic Mapping Study. arXiv:2212.10693 [cs.SE]

[2] Bakhtiyar Ahmed, Thomas Dannhauser, and Nada Philip. 2018. A Lean Design Thinking Methodology (LDTM) for Machine Learning and Modern Data Projects. In *2018 10th Computer Science and Electronic Engineering (CEEC)*. 11–14. https://doi.org/10.1109/CEEC.2018.8674234

[3] Alhejab Alhazmi and Shihong Huang. 2020. Integrating Design Thinking into Scrum Framework in the Context of Requirements Engineering Management. In *Proceedings of the 3rd International Conference on Computer Science and Software Engineering*. 33–45. https://doi.org/10.1145/3403746.3403902

[4] Silvio Alonso, Antonio Pedro Santos Alves, Lucas Romao, Hélio Lopes, and Marcos Kalinowski. 2025. Define-ML: An Approach to Ideate Machine Learning-Enabled Systems. In *2025 51st Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*. Springer.

[5] Antonio Pedro Santos Alves, Marcos Kalinowski, Görkem Giray, Daniel Mendez, Niklas Lavesson, Kelly Azevedo, Hugo Villamizar, Tatiana Escovedo, Helio Lopes, Stefan Biffl, Jürgen Musil, Michael Felderer, Stefan Wagner, Teresa Baldassarre, and Tony Gorschek. 2024. Status Quo and Problems of Requirements Engineering for Machine Learning: Results from an International Survey. In *Product-Focused Software Process Improvement*, Regine Kadgien, Andreas Jedlitschka, Andrea Janes, Valentina Lenarduzzi, and Xiaozhou Li (Eds.). Springer Nature Switzerland, Cham, 159–174.

[6] S. S. Anand, A. R. Patrick, J. G. Hughes, and D. A. Bell. 1998. A data mining methodology for cross-sales. *Know.-Based Syst.* 10, 7 (May 1998), 449–461. https://doi.org/10.1016/S0950-7051(98)00035-5

[7] Atlas.TI 2024. https://atlasti.com.

[8] V.R. Basili and H.D. Rombach. 1988. The TAME project: towards improvement-oriented software environments. *IEEE Transactions on Software Engineering* 14, 6 (June 1988), 758–773. https://doi.org/10.1109/32.6156

[9] Peter Cabena, Pablo Hadjinian, Rolf Stadler, Jaap Verhees, and Alessandro Zanasi. 1998. *Discovering data mining: from concept to implementation.* Prentice-Hall, Inc., USA.

[10] J. Corbin and A. Strauss. 2014. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (4th ed.). SAGE Publications. https://books.google.com.br/books?id=hZ6kBQAAQBAJ

[11] CRISP-DM - version 2.0 2008. CRISP-DM - CRoss-Industry Standard Process for DM. (version 1.0 of 1999).

[12] Kevin Dewalt. 2017. The AI Canvas - The strategic framework for enterprise deep learning. https://medium.com/the-business-of-ai/the-ai-canvas-7a8717cddbe9

[13] Louis Dorard. 2015. Machine Learning Canvas. https://www.machinelearningcanvas.com/

[14] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. 1996. Knowledge discovery and data mining: towards a unifying framework. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (Portland, Oregon) *(KDD'96)*. AAAI Press, 82–88.

[15] Raphael Augusto Parreiras Gallegos. 2023. *Ferramentas de Gestão Voltadas para Melhoria da Qualidade nas Empresas.* Freitas Bastos Editora.

[16] Görkem Giray. 2021. A software engineering perspective on engineering machine learning systems: State of the art and challenges. *Journal of Systems and Software* 180 (2021), 111031.

[17] John W Goodell, Satish Kumar, Weng Marc Lim, and Debidutta Pattnaik. 2021. Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance* 32 (2021), 100577.

[18] Daniel Gustafsson. 2019. Analysing the Double diamond design process through research & implementation. https://api.semanticscholar.org/CorpusID:199016905

[19] Khan Mohammad Habibullah, Gregory Gay, and Jennifer Horkoff. 2023. Non-functional requirements for machine learning: understanding current use and challenges among practitioners. *Requirements Engineering* 28, 2 (Jan. 2023), 283–316. https://doi.org/10.1007/s00766-022-00395-3

[20] Jennifer Hehn et al. 2020. On Integrating Design Thinking for Human-Centered Requirements Engineering. *IEEE Software* 37, 2 (2020), 25–31. https://doi.org/10.1109/MS.2019.2957715

[21] Jennifer Hehn and Falk Uebernickel. 2018. The Use of Design Thinking for Requirements Engineering: An Ongoing Case Study in the Field of Innovative Software-Intensive Systems. In *2018 IEEE 26th International Requirements Engineering Conference (RE)*. 400–405. https://doi.org/10.1109/RE.2018.00-18

[22] Jennifer Hehn, Falk Uebernickel, and Daniel Mendez Fernandez. 2018. DT4RE: Design Thinking for Requirements Engineering: A Tutorial on Human-Centered and Structured Requirements Elicitation. In *2018 IEEE 26th International Requirements Engineering Conference (RE)*. 504–505. https://doi.org/10.1109/RE.2018.00074

[23] Hans-Martin Heyn, Eric Knauss, Amna Pir Muhammad, Olof Eriksson, Jennifer Linder, Padmini Subbiah, Shameer Kumar Pradhan, and Sagar Tungal. 2021. Requirement Engineering Challenges for AI-intense Systems Development. arXiv:2103.10270 [cs.LG]

[24] Boyue Caroline Hu, Rick Salay, Krzysztof Czarnecki, Mona Rahimi, Gehan Selim, and Marsha Chechik. 2020. Towards requirements specification for machine-learned perception based on human performance. In *2020 IEEE Seventh International Workshop on Artificial Intelligence for Requirements Engineering (AIRE)*. IEEE, 48–51.

[25] Fuyuki Ishikawa and Nobukazu Yoshioka. 2019. How Do Engineers Perceive Difficulties in Engineering of Machine-Learning Systems? - Questionnaire Survey. In *2019 IEEE/ACM Joint 7th International Workshop on Conducting Empirical Studies in Industry (CESI) and 6th International Workshop on Software Engineering Research and Industrial Practice (SERIP)*. 2–9. https://doi.org/10.1109/CESSER-IP.2019.00009

[26] ISO. 2017. ISO/IEC/IEEE International Standard - Systems and Software Engineering – Software Life Cycle Processes. , 157 pages. https://doi.org/10.1109/IEEESTD.2017.8100771

[27] Fei Jiang, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Sufeng Ma, Yilong Wang, Qiang Dong, Haipeng Shen, and Yongjun Wang. 2017. Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology* 2, 4 (2017).

[28] Danijel Kučak, Vedran Juričić, and Goran Đambić. 2018. MACHINE LEARNING IN EDUCATION-A SURVEY OF CURRENT RESEARCH TRENDS. *Annals of DAAAM & Proceedings* 29 (2018).

[29] Oscar Marban, Gonzalo Mariscal, and Javier Segovia. 2009. A Data Mining Knowledge Discovery Process Model. In *Data Mining and Knowledge Discovery in Real Life Applications*, Julio Ponce and Adem Karahoca (Eds.). IntechOpen, Rijeka, Chapter 1. https://doi.org/10.5772/6438

[30] Mariana Martins, Taciana Kudo, and Renato Bulcão-Neto. 2024. A Qualitative Study on Requirements Engineering Practices in an Artificial Intelligence Unit of the Brazilian Industrial Research and Innovation Company. In *Anais do XXVII Congresso Ibero-Americano em Engenharia de Software* (Curitiba/PR). SBC, Porto Alegre, RS, Brasil, 46–60. https://sol.sbc.org.br/index.php/cibse/article/view/28438

[31] Mariana Martins, Taciana Kudo, and Renato Bulcão-Neto. 2024. A Requirements Engineering Process for Machine Learning Innovation Projects. In *Anais do XXVII Workshop de Engenharia de Requisitos* (Buenos Aires/ARG). Even3, Brazil, Buenos Aires, Argentina. http://wer.inf.puc-rio.br/WERpapers/

[32] Mariana Crisostomo Martins, Lívia Mancine C. Campos, João Lucas R. Soares, Taciana Novo Kudo, and Renato F. Bulcão-Neto. 2025. Requirements Engineering for Machine Learning-Based AI Systems: A Tertiary Study. *Journal of Software Engineering Research and Development* 13, 2 (Sep. 2025), 13:129 – 13:142. https://doi.org/10.5753/jserd.2025.4892

[33] Elizamary Nascimento, Anh Nguyen-Duc, Ingrid Sundbø, and Tayana Conte. 2020. Software engineering for artificial intelligence and machine learning software: A systematic literature review. arXiv:2011.03751 [cs.SE] https://arxiv.org/abs/2011.03751

[34] Lauriane Pereira et al. 2021. Towards an understanding of benefits and challenges in the use of design thinking in requirements engineering. 1338–1345. https://doi.org/10.1145/3412841.3442008

[35] Carlos E. Perez. 2017. *The Deep Learning Ai Playbook: Strategy for Disruptive Artificial Intelligence.* Intuition Machine.

[36] S.L Pfleeger. [n. d.]. Understanding and improving technology transfer in software engineering. *Journal of Systems and Software* 47, 2 ([n. d.]), 111–124. https://doi.org/10.1016/S0164-1212(99)00031-X

[37] Lucas Romao, Marcos Kalinowski, Clarissa Barbosa, Allysson Allex Araújo, Simone D. J. Barbosa, and Helio Lopes. 2024. Agile Minds, Innovative Solutions, and Industry-Academia Collaboration: Lean RD Meets Problem-Based Learning in Software Engineering Education. arXiv:2407.15982 [cs.SE] https://arxiv.org/abs/2407.15982

[38] Per Runeson and Martin Höst. 2008. Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering* 14, 2 (Dec. 2008), 131–164. https://doi.org/10.1007/s10664-008-9102-8

[39] Viswanath Venkatesh and Hillol Bala. 2008. Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences* 39, 2 (May 2008), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

[40] Hugo Villamizar, Tatiana Escovedo, and Marcos Kalinowski. 2021. Requirements engineering for machine learning: A systematic mapping study. In *2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*. IEEE, 29–36.

[41] Hugo Villamizar, Marcos Kalinowski, and Helio Lopes. 2022. Towards Perspective-Based Specification of Machine Learning-Enabled Systems. arXiv:2206.09760 [cs.SE]

[42] Hugo Villamizar, Marcos Kalinowski, Hélio Lopes, and Daniel Mendez. 2024. Identifying concerns when specifying machine learning-enabled systems: A perspective-based approach. *Journal of Systems and Software* 213 (2024), 112053. https://doi.org/10.1016/j.jss.2024.112053

[43] Hugo Villamizar, Marcos Kalinowski, Marx Viana, and Daniel Méndez Fernández. 2018. A Systematic Mapping Study on Security in Agile Requirements Engineering. In *2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*. 454–461. https://doi.org/10.1109/SEAA.2018.00080

[44] Andreas Vogelsang and Markus Borg. 2019. Requirements Engineering for Machine Learning: Perspectives from Data Scientists. arXiv:1908.04674 [cs.LG]

[45] Zhiyuan Wan, Xin Xia, David Lo, and Gail C. Murphy. 2020. How does Machine Learning Change Software Development Practices? *IEEE Transactions on Software Engineering* (2020), 1–1. https://doi.org/10.1109/tse.2019.2937083

[46] Matei Zaharia, Andrew Chen, Aaron Davidson, Ali Ghodsi, Sue Ann Hong, Andy Konwinski, Siddharth Murching, Tomas Nykodym, Paul Ogilvie, Mani Parkhe, et al. 2018. Accelerating the machine learning lifecycle with MLflow. *IEEE Data Eng. Bull.* 41, 4 (2018), 39–45.