QUESTRL: A Q&A Framework for Designing Trustworthy Reinforcement Learning Systems

Katherine R. Dearstyne
University of Notre Dame
Notre Dame, IN, USA
kdearsty@nd.edu

Pedro (Tony) Alarcon Granadeno,
Theodore Chambers
University of Notre Dame
Notre Dame, IN, USA
{palarcon, tchambe2}@nd.edu

Jane Cleland-Huang
University of Notre Dame
Notre Dame, IN, USA
JaneHuang@nd.edu

***Abstract*—Cyber-Physical Systems (CPS) increasingly leverage Reinforcement Learning (RL) to adapt dynamically to changing environments and optimize performance over time. While RL enhances efficiency and safety by enabling autonomous adjustments to unexpected conditions and hazard avoidance, it also introduces significant risks, as learned behaviors may lead to unpredictable or unsafe actions in real-world deployment. Therefore, integrating risk management into RL system design is essential. In this paper, we propose the *QuestRL Framework*, a question-driven approach that translates high-level safety guidelines into RL-specific considerations. This framework helps RL practitioners address key risks early in development, informing new or existing system requirements while ensuring traceability to risk management objectives. To evaluate its effectiveness, we conducted a study across two use cases, engaging six RL experts in developing system requirements with and without the framework. Our findings suggest that the framework promotes critical thinking and helps practitioners identify additional risk factors, ultimately supporting safer RL deployment.**

***Index Terms*—reinforcement learning, cyber-physical systems, safety**

I. Introduction

To achieve higher levels of autonomy, Cyber-Physical Systems (CPS) often leverage Reinforcement Learning (RL)—a machine learning (ML) paradigm where agents learn optimal decisions through continuous interaction with their environment, guided by rewards and penalties [2]. Across domains such as autonomous vehicles [3], smart grids [4], and Small Uncrewed Aerial Systems (sUAS) [5]–[7], RL helps enable adaptive behaviors in dynamic and uncertain environments. However, it also introduces unique verification challenges compared to traditional ML, as RL agents must operate in inherently unpredictable settings that evolve in complex and unforeseen ways—often as a result of the agents' own actions [2]. Factors such as sparse and delayed reward signals, partial observability, multi-agent interactions, and potential continuous learning during deployment further complicate prediction and verification. These challenges become more pronounced in safety-critical applications, where ensuring reliability under all conditions becomes both more challenging and consequential. The compounded uncertainty of RL and CPS fundamentally complicates the development of robust safety cases [8], making it essential to integrate risk management strategies throughout the life cycle of RL-based systems.

To address these risks, standards and frameworks, such as ISO/IEC 23894:2023 [9] and the NIST AI Risk Management Framework (AI RMF) [10], provide essential guidance on AI safety. However, they remain broad in scope and lack specific considerations tailored to the unique challenges of RL. Practitioners are often left to interpret these high-level principles without clear direction on how to translate them into concrete, implementable strategies for their domain. As a result, defining comprehensive requirements demands expertise not only in RL but also in system safety and requirements engineering. These knowledge gaps, combined with the significant time investment required, often lead to requirements specifications being ad hoc—or neglected entirely [11], [12], delaying risk mitigation and undermining safety assurance [13], [14]. Without a structured approach to integrating safety considerations from the outset, RL-based systems risk a cycle of retroactive fixes rather than proactive design.

To address this challenge, we introduce the *QuestRL Framework*, a **question**-driven method that provides a structured approach to the design and specification of **RL**-based systems. Rather than leaving practitioners to interpret broad safety guidelines on their own, *QuestRL* deconstructs high-level objectives, derived from existing standards, into concrete, targeted questions that guide RL engineers in documenting critical design decisions for trustworthy RL in CPS. These questions are deliberately designed to encompass general AI safety principles while also addressing the unique challenges of RL. This combined focus helps tackle the harder problems of RL verification without overlooking foundational ML safety concerns. *QuestRL* also builds on a previously validated process to reverse engineer requirements from these design decisions [15], [16], enabling the creation of standalone RL specifications or integration into broader system requirements. Together, these elements enable the systematic consideration of safety and trustworthiness early in the development process.

The primary contributions of this paper are:

- The *QuestRL Framework*, which maps AI-related objectives from established standards into targeted ques-

Supplementary materials available at https://zenodo.org/records/15825459 [1].

TABLE I: AI System Objectives with descriptions, example ‘ASKS’, and the number of high and low ASKs for each.

Objective	Description	Example ASK	# of ASKs (high; low)
Accountability (System)	Provides traceability of decisions and actions back to responsible entity.	Does the system enable detailed post-mortem analysis, including replaying scenarios to trace the reasoning behind decisions?	1; 1
AI Expertise	Ensures users can understand system behavior and intervene if errors occur.	How will the system demonstrate examples of expected, degraded, and failure behaviors in a way that is interpretable for the overseer?	1; 2
Availability & Quality of Training and Test Data	Ensures data is current, relevant, and appropriately diverse.	How are high-risk or low-probability scenarios identified and prioritized into testing and replay strategies?	4; 4
Fairness	Prevents biased outcomes from objectives, data, or human feedback.	What methods/metrics are used to identify biases in the agent’s decision-making processes?	2; 2
Maintainability	Enables modification to correct defects or adapt to new requirements.	What mechanisms are in place to detect when performance/behaviors degrade to a point that requires policy updates or retraining?	1; 1
Privacy	Protects individuals’ control over their personal data during collection, storage, and processing.	What safeguards prevent data leakage or misuse during model training and inference?	1; 3
Robustness	Ensures stable performance under diverse and uncertain conditions.	How does the system infer missing or uncertain state information when observations are incomplete or unreliable?	10; 11
Accuracy	Ensures results are consistent with true values or accepted standards.	Could the current reward design inadvertently incentivize undesired behaviors? What measures are in place to identify and mitigate such risks?	4; 7
Verifiability	Ensures system correctness and safety under defined conditions.	Are formal verification methods used to ensure the correctness and safety of the RL system? If so, which specific methods are applied?	1; 0
Safety	Prevents the system from endangering life, health, property, or the environment.	How are the constraints or penalties balanced with the reward so the agent is not tempted toward a higher reward at the cost of safety?	5; 10
Security	Protects the system from data poisoning, adversarial attacks, and model theft.	What fallback systems, predefined controllers, or recovery strategies are implemented to address situations where the RL agent fails, behaves unpredictably, or enters an unsafe state?	1; 1
Transparency	Provides stakeholders with clear explanations of system outputs and decisions.	What conditions or thresholds will trigger alerts, and how will the system communicate these to the overseer?	4; 4

tions addressing both foundational ML concerns and RL-specific challenges in CPS.

- An automated method for generating RL-specific requirements from design decisions elicited through *QuestRL*.
- Initial validation of the *QuestRL Framework* in two CPS scenarios, demonstrating its effectiveness in promoting safety and trustworthiness considerations.
- Discussion of potential extensions to *QuestRL*, including broadening its applicability to other AI/ML paradigms.

The remainder of the paper is structured as follows: Section II presents the *QuestRL Framework* and its role in AI risk management. Section III outlines the evaluation methodology, followed by results in Section IV. Sections V to VII cover related work, threats to validity, conclusions and future work.

II. The *QuestRL Framework*

The *QuestRL Framework* adopts a *Question and Answer* approach to guide RL Engineer in considering critical design decisions that might otherwise be deferred until later in the process. We opted for a Q&A format rather than a checklist to encourage design-thinking and prompt users to document their design decisions for later use in generating a requirements specification. Further, there are always trade-offs in designing a framework, especially in balancing the level of abstraction versus specificity. Our aim was to offer enough detail to prompt meaningful reflection without dictating specific choices, allowing practitioners to explore diverse, context-appropriate solutions.

QuestRL is formulated as a hierarchy of nodes, composed of four key components. **Objectives**, drawn primarily from standards and white papers, represent high-level goals for risk mitigation in RL-based systems. These are refined into **ASKs** which pose specific questions addressing AI safety considerations tailored to the unique challenges of RL. Additionally, the framework incorporates **Context** questions which help users define system scope and document environmental assumption before beginning the process. Finally, **Requirements** nodes are generated automatically from responses to the ASKs. Users can accept, reject, and modify the generated requirements as appropriate, ensuring practitioners retain control over the final decisions. Each generated requirement includes a system-level requirement along with a set of supporting design decisions. These design decisions may be presented as bullet points within the requirement or as separate node types, serving as the leaves of the hierarchy.

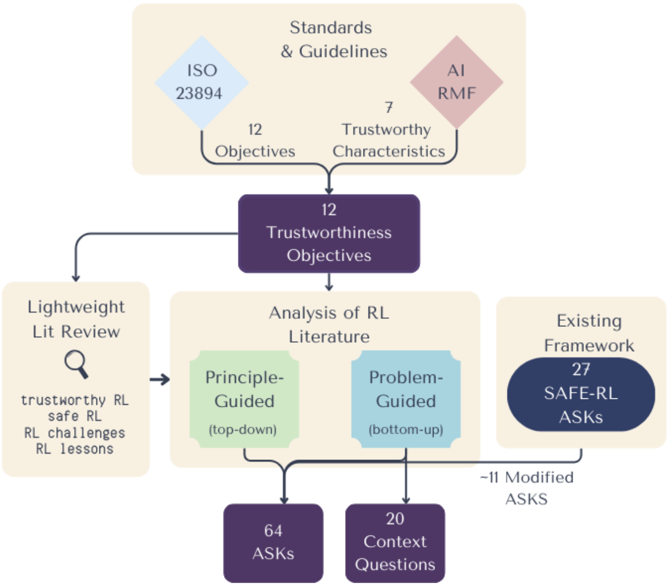


Fig. 1: Depicts the development of *QuestRL*, starting from 2 standards to derive 12 (Objectives). A lit review then informed 64 ASKs and 20 context questions. 11 ASKs were adapted from the SAFE-RL framework.

A. *Components of QuestRL Framework*

*1) **Objectives**:* Objectives provide the overall structure of the framework. As depicted in Figure 1, they are primarily derived from ISO 23894 [9] and the NIST AI RMF [10]. While ISO 23894 outlines 12 objectives and NIST identifies 7 trustworthiness characteristics, significant overlap allowed us to merge them into a unified set closely aligned with ISO 23894, with several key adjustments.

First, we exclude *environmental impact* as it pertains to higher-level systemic concerns beyond the technical scope of RL. Second, we narrowed the scope of *accountability* to focus on *system-level* accountability rather than organizational. Additionally, ISO 23894 lists *robustness* as a standalone objective, whereas the AI RMF groups *robustness*, *accuracy*, and *reliability* under the objective of "valid and reliable." To capture these distinctions more effectively, we include them in *QuestRL* as three as separate objectives. These modification lead to the 12 objectives described in Table I.

*2) **ASKs**:* Question nodes, referred to as 'ASKs', are a critical component of the *QuestRL Framework*. Unlike checklists or Safety Assurance Cases, which may reinforce bias by affirming claims [17], `ASK` nodes prompt stakeholders to proactively consider how their design can address each objective, thereby increasing the safety of the delivered system.

ASKs were identified from two primary sources. The first source was a review of the existing literature. To conduct this review, we performed multiple targeted searches combining 'Reinforcement Learning' with terms such as 'Safety', 'Trustworthy', 'Challenges', and 'Lessons' in Google Scholar. This yielded approximately 3.39 million results across all queries. For each query, we sorted the papers by relevance score and analyzed the top 20. To ensure broad applicability to RL in CPS, we excluded papers that focused on specific RL algorithms or systems. Additionally, we prioritized papers published within the last five years. For those that met our criteria, we also examined cited upstream works to capture relevant foundational contributions. The final selection of 10 papers is presented in Table II.

During our literature review we adopted a hybrid approach that included (a) a *principle-driven*, top-down approach that derived ASKs from examples that successfully addressed the framework's objectives, and (b) *a problem-driven*, bottom-up approach using thematic analysis to identify recurring concerns, which were then mapped to relevant objectives. When a theme aligned with multiple objectives, we assigned it accordingly. Each identified theme was then formulated into a structured question aimed at fostering critical reflection on how the system could satisfy the corresponding objective.

TABLE II: Key papers from the lightweight literature review.

Challenges of Real-World Reinforcement Learning: Definitions, Benchmarks and Analysis	[18]
How Reinforcement Learning Systems Fail and What to do About It	[19]
Using Deep Reinforcement Learning And Formal Verification in Safety Critical Systems: Strategies and Challenges	[20]
Challenges in the Verification of Reinforcement Learning Algorithms	[8]
Trustworthy Reinforcement Learning Against Intrinsic Vulnerabilities: Robustness, Safety, and Generalizability	[21]
Safe and Robust Reinforcement Learning: Principles and Practice	[22]
How to Train Your Robot with Deep Reinforcement Learning; Lessons We've Learned	[23]
Common challenges of deep reinforcement learning applications development: an empirical study	[24]
A Review of Safe Reinforcement Learning: Methods, Theories, and Applications	[25]
Multi-Agent Reinforcement Learning: Methods, Applications, Visionary Prospects, and Challenges	[26]

The second source of information for establishing ASKs was an existing Q&A approach [27] which was developed through an iterative design process with feedback from RL engineers. However, this approach neither incorporated existing AI-related standards nor underwent formal evaluation. Nonetheless, it included 27 questions, 11 of which were relevant to the standards-driven objectives in *QuestRL*. These questions were refined to achieve a higher level of specificity and integrated into *QuestRL*.

This process resulted in 27 high-level and 37 lower-level ASKs. An example ASK for each objective is presented in Table I, along with the total number of high and low level ASKs associated with the objective. Because some ASKs map to multiple objectives, they are counted in each relevant category. We include an example of a high-level ASK, broken down into two low-level, specific ASKs in Figure 2.

In line with our philosophy of fostering critical reflection rather than prescribing specific solutions, *QuestRL* is designed to be adaptable. Additional ASKs can be introduced, and existing ones can be tailored to suit specific projects or contexts, evolving alongside advancements in RL.

*3) **Context**:* Establishing a clear set of assumptions about the physical and operational environment, adjacent system behavior, users, and the development process is essential for guiding design decisions and specifying requirements [28]. To support this, *QuestRL* includes a set of questions aimed at identifying RL-related assumptions and contextual information. These questions are distinct from ASKs because they do not target high-level objectives directly; instead, they are intended to be answered beforehand, ensuring designers explicitly consider assumptions when making critical decisions.

Prior research has identified core categories of assumptions within RL [8], [29]. We incorporate several of these categories as the foundation for our assumption framework. We also add a category for users because of the emphasis placed on end-user considerations by both the ISO 23894 and AI RMF. Each of these categories is associated with several targeted questions designed to prompt practitioners to reflect on the underlying assumptions and context surrounding their system. As with the ASKs, these questions were informed by literature (e.g., [30]–[34]. Descriptions of each category along with examples is provided in Table III.

TABLE III: The 5 context categories with example questions.

Context Category	Example Question
Scope	What are the primary **objectives** that the RL agent should accomplish?
Users	What is the level of **AI expertise** of the user(s)?
Learning	Is learning **online** (adaption during interactions with environment) or **offline** (learns solely from pre-collected dataset)?
Environment	Is the environment **fully** (agent has access to complete information about the current state) or **partially** observable (agent only receives partial or noisy information about the current state)?
Platform	What is the available **processing power** for the RL agent on the device?

*4) **Requirements**:* While ASKs play a fundamental role in guiding design thinking, they—along with the answers provided by users—also serve as the primary source of information for deriving candidate system requirements. Given a set of ASKs for which design decisions have already been made, *QuestRL* leverages a Large Language Model (LLM) to generate a corresponding set of candidate requirements. For this study, we employed GPT-4 due to its demonstrated effectiveness in requirements engineering tasks [35], though future work should explore the impact of using other LLMs within the framework. Our approach builds upon previously established and validated techniques for generating requirements from design artifacts [15], [36].

ASK 3 **How is accuracy assessed in the system?**
- **ASK 3.1** What metrics are used to evaluate performance, and how are acceptable ranges determined?
- **ASK 3.2** Are there validation procedures to compare RL-generated actions with those of expert human operators or traditional controllers?

Fig. 2: Example high-level ASK with low-level ASKs

To generate the requirements, we follow the process illustrated in Figure 3. We constructed a prompt, as shown in Figure 4, that includes all ASKs under each objective, the associated responses provided by the user, and any preexisting requirements. The LLM was instructed to output its response in JSON format, with fields for the requirement title, body, rationale, associated design decisions, and a list of ASKs it relied upon when generating the requirement. Crucially, the LLM was prompted to rely solely on the information provided in the ASKs, user responses, and existing requirements to avoid introducing any fabricated content.

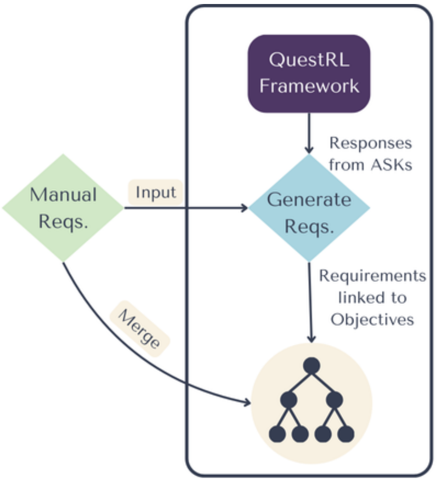


Fig. 3: Demonstrates applying *QuestRL*, where the ASK responses from the user are used as input to generate requirements. If manual requirements already exist, they can be merged with the output of *QuestRL*.

Although expert review remains necessary to ensure alignment with design intent and objectives, this process offers several advantages. First, it produces structured, traceable specifications that clarify system expectations and support verification. Second, it enhances change management by serving as a stable reference point that abstracts implementation details, facilitating impact analysis as design decisions evolve. Additionally, by grounding requirement generation in user responses, the process ensures specifications reflect actual design intent and maintain traceability to stated objectives. These generated requirements can serve as standalone RL specifications or be integrated with existing requirements into a system-level specification.

B. QuestRL within AI RMF Core

QuestRL is designed to complement existing AI risk management frameworks. Here we show how it integrates within the four core functions of the AI RMF: GOVERN, MAP, MEASURE, and MANAGE [10].

Prompt:
# TASKS:
You are an expert engineer of deep-RL systems. You are building a RL system. Certain design decisions about building the system have been made and are described in the form of Q&A below. Each question/answer is related to an overarching objective described in the AI Risk Management ISO. The specific objective currently being focused on is *Maintainability*. Your goal is to create a set of requirements that describe the design decisions in the answers. These requirements, if satisfied, will fulfill the given objective. Currently, some requirements have already been created. You will be revising these and creating any necessary new ones to cover all the information given by the ASKs. To create and refine the requirement set, go through the user responses. For each of the responses, identify if there any new details or information not found in the original requirements. If so, decide whether the information should either supplement an existing requirement or if a new requirement should be created. Then, output both the revised and newly created requirements. Format your Requirements as JSON as described below:
Note: The response format has been omitted for brevity due to space constraints.
# ORIGINAL REQUIREMENTS
`[REQ1]` The traffic signal control system shall provide detailed logs and analytics for monitoring performance and debugging issues.

  * Implement a log aggregation system that collects data from all sensors and devices.
  * Develop an analytics platform that provides insights into system performance, including queue lengths, vehicle speeds, and conflict rates.

Continues... (Omitted for brevity)
# Q & A for Design Decisions

  * What mechanisms are in place to detect when performance/behaviors degrade to a point that requires policy updates or retraining?
    ** The primary indicator of safety degradation is constrain violation rate. If the rate of CVR increases beyond critical limit or consistently goes beyond threshold
    ** Detection of reward trend deviation such as continuous downward trend or rolling average drops greater than threshold over a certain period of time.
    ** At this point, safe fallback activation can be triggered

Continues... (Omitted for brevity)

Response: {
"Title": "Performance Degradation Detection",
"SystemRequirement": "The system must detect when performance degrades and requires policy updates.",
"ImplementationDetails": "The primary indicator of safety degradation is an increase in Constraint Violation Rate (CVR), Detect reward trend deviations such as continuous downward trends or rolling average drops beyond a threshold over time, Activate a safe fallback mechanism when degradation is detected",
"Rationale": "Detecting performance degradation ensures timely interventions to maintain safety."
}
Continues... (Omitted for brevity)

Fig. 4: Shows the prompt used for requirements generation with Maintainability as an example and a subset of P2's requirements and ASKs. One example of the LLM-generated requirements for P2 is also shown beneath the prompt.

The **GOVERN** function ensures that "risks and potential impacts are identified, measured, and managed effectively and consistently." *QuestRL* reinforces GOVERN 1.2's requirement to integrate trustworthy AI characteristics into organizational processes by establishing traceability from high-level objectives to specific system requirements and design decisions. For the **MAP** function, which establishes the context to frame risks related to an AI system," the *QuestRL Framework* employs structured questions to document system assumptions and contextual factors specific to the RL component. These ASK responses inform requirement generation, directly supporting MAP 1.7's directive that "system requirements are elicited and understood from stakeholders." The **MEASURE** function "employs quantitative, qualitative, or mixed-method tools, techniques, and methodologies to analyze, assess, benchmark, and monitor AI risk." *QuestRL* supports this through targeted ASKs that elicit evaluation details, such as ASK 3 shown in Figure 2. While the *QuestRL Framework* does not directly implement the **MANAGE** function—defined as "allocating risk management resources to mapped and measured risks"—it facilitates this process by documenting mitigation strategies that support risk prioritization and analysis.

Across all four functions, the *QuestRL Framework* serves as a supporting tool rather than a replacement. It equips practitioners and stakeholders with critical insights into risk identification, mitigation strategies, evaluation methods, and traceability to better enact the AI RMF within the RL context.

III. Evaluation

To evaluate the *QuestRL Framework* we posed three research questions as follows:

- RQ1: To what extent does the *QuestRL Framework* aid developers in specifying a comprehensive set of requirements for safe and trustworthy RL-based systems?
- RQ2: To what extent are requirements generated by *QuestRL* acceptable to RL-Engineers?
- RQ3: How do developers perceive *QuestRL* compared to manually creating requirements from scratch?

To address these questions we conducted an IRB approved study in which six RL experts were each tasked with applying the framework to the design and specification of a Cyber-Physical System with respect to specific *QuestRL* objectives.

A. Participants

We recruited six participants from Upwork, a freelancing platform that connects professionals with clients. Upwork was chosen due to its large pool of technical consultants. We limited the study to six participants due to the cost of hiring skilled RL developers for a task that we estimated would take up to 10 hours. To identify suitable candidates, we only invited freelancers who had explicitly described themselves as RL experts in their profiles to apply. As part of the application process, candidates completed a pre-survey in which they detailed prior RL projects and answered RL-specific questions. We excluded anyone who had not completed at least one RL project. We selected the first 6 applicants that met these criteria. Among the six participants, half had experience with industrial RL applications, primarily in robotics, while the other half had conducted RL-focused research as Ph.D. students. Compensation was based on their Upwork hourly rate ($18–35), with a 10-hour cap to stay within budget. All participants were able to complete the assigned tasks within the 10-hour time limit.

B. Use Cases

Our study focused on two distinct use cases, each derived from a research paper applying RL to CPS. The first (UC1) focused on optimizing a traffic signal control (TSC) at a four-way intersection, inspired by Zhang et al. [37]. The goal of the system was to dynamically adjust traffic signal phases to minimize congestion, delays, and conflicts. The second (UC2), based on Smith et al. [38], applied RL to enable a quadruped robot to navigate hazardous and unpredictable environments. The goal was for the robot to autonomously adapt its movements to environmental conditions, allowing it to traverse uneven, deformable, or obstructed terrain without relying on predefined gaits.

Participants were given a one-page problem description detailing the deployment environment, available sensor information, and the hardware used. As participants in our study were previously unfamiliar with these systems, we also prefilled the *context* nodes with information provided in the corresponding papers. Participants were instructed to incorporate this context and assumptions when responding to the ASKs.

C. Tasks

Three participants (P1, P2, and P3) were assigned to UC1 and three (P4, P5, P6) to UC2. Due to the volume and depth of the questions, we divided the 12 objectives into three groups, with each participant addressing four objectives from their assigned use case:

- Group 1: Robustness, accuracy, verifiability, and availability & quality of training and test data *(assigned to P1, P4).*
- Group 2: Safety, security, maintainability, and privacy *(assigned to P2, P5).*
- Group 3: Transparency, accountability, AI expertise, and fairness *(assigned to P3, P6).*

Participants began by reviewing their use case and assigned objectives, then completed three tasks over 3–7 days. Tasks were completed sequentially, and each participant only received instructions for the next task after the previous one was submitted. On average, participants reported taking 3 hours for Tasks 1-2 and 1 hour for Task 3.

Task 1 **Define System Requirements (Baseline):** To establish a baseline for how each participant would manually construct requirements without *QuestRL*, each participant identified key system requirements necessary to achieve their assigned objectives. They were instructed to specify requirements using the EARS event-driven template [39] and to attach a bulleted list of possible design solutions.

Task 2 **Respond to Objective ASKs:** After specifying their requirements, participants were given an Excel sheet containing ASKs related to their assigned objectives. We opted to use an excel sheet for accessibility reasons as our RL-Engineers were working remotely. They answered all relevant ASKs and provided justification for any that they left unanswered. For example, questions related to multi-agent solutions were not relevant in a single-agent scenario.

Task 3 **Review Generated Requirements:** Finally, using each participant's ASK responses and the prompt shown in Figure 4, we leveraged GPT-4 [40] to generate a personalized set of requirements for each participant. As elaborated in Section II-A4, the prompt specifically incorporated each participant's responses to the ASKs to ensure the generated requirements reflected their individual input. From each participant's generated set, we randomly selected four requirements that included at least one from each of their assigned objectives and asked the participant to review these requirements. Participants rated each requirement using the following rubric:

i. **Reject** – This requirement is irrelevant, unnecessary, or incorrect.
ii. Accept with **Major** Modifications – The requirement is somewhat relevant but would need significant changes.
iii. Accept with **Minor** Modifications – The requirement is useful but requires slight revisions.
iv. Accept **As Written** - The requirement is well-written, relevant, and can be added without changes.

For rejected or modified requirements, participants were asked to explain their reasoning and to describe any necessary changes. We include one of the selected requirements per participant in Figure 5.

D. Feedback

At the conclusion of the study, all participants completed an exit questionnaire consisting of seven questions focused on their experience with *QuestRL* and its associated tasks:

1) Were there any additional questions related to your assigned objective(s) that you believe should have been included in the ASKs from Task 2? If yes, please specify.
2) What challenges or difficulties did you encounter throughout Tasks 1 and 2?
3) If you were to revisit your initial system requirements (Task 1) after completing Task 2, would you make any revisions? If so, what high-level changes would you consider? (Note: You are not required to make these changes—just summarize your thoughts.)
4) Are there any improvements that you would suggest to enhance the *QuestRL* process in the future?
5) Did you find Task 1 or Task 2 to be more challenging? Explain why.

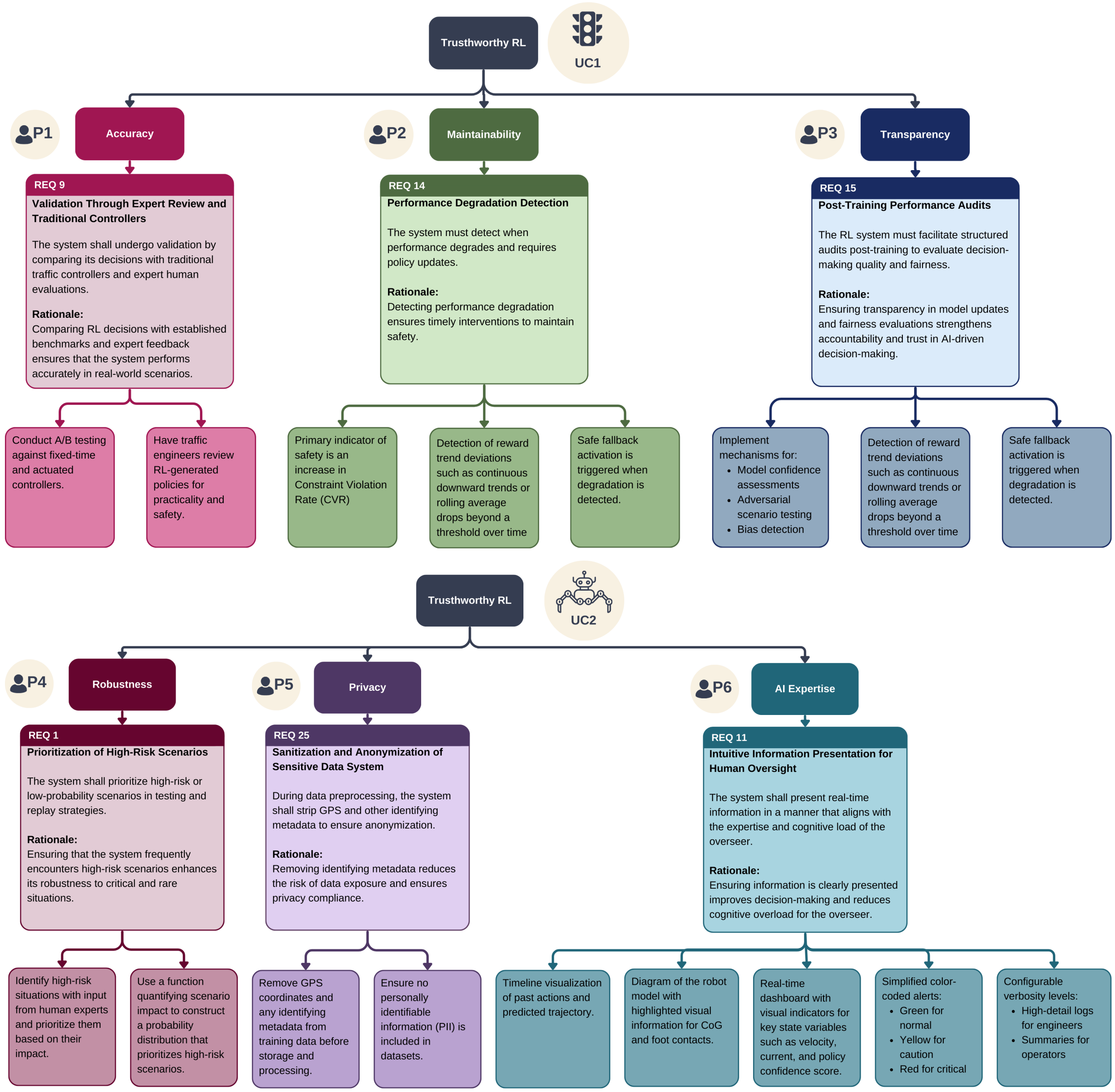
6) To what extent did you rely on external sources to answer the questions in Task 2? For each of the following, select one of the options: [Not at all, Occasionally, Frequently, Heavily]
- ChatGPT or other Large Language Models (LLMs)
- Google or other search engines
- Academic papers or technical documentation

IV. Results

In this section we analyze the collected information from the study to answer each of the three research questions.

A. *RQ1: To what extent does the QuestRL Framework aid developers in specifying a comprehensive set of requirements for safe and trustworthy Reinforcement Learning-based systems?*

To answer this question, we conducted a quantitative analysis of the coverage of the *QuestRL Framework*'s objectives in each participant's initial specification (Task 1), versus the coverage provided by the generated requirements. Results are reported in Table IV for each participant. The first row labeled 'Original' shows the number of requirements that each participant authored in their original specification; the 'Generated' row shows the number of relevant, distinct requirements generated by the LLM. These are further divided into 'new' (not present in the original set) and 'modified' (additional details added to an existing requirement). The results indicate



Trusthworthy RL
UC1
P1
Accuracy
REQ 9
Validation Through Expert Review and Traditional Controllers
The system shall undergo validation by comparing its decisions with traditional traffic controllers and expert human evaluations.
Rationale:
Comparing RL decisions with established benchmarks and expert feedback ensures that the system performs accurately in real-world scenarios.
Conduct A/B testing against fixed-time and actuated controllers.
Have traffic engineers review RL-generated policies for practicality and safety.
P2
Maintainability
REQ 14
Performance Degradation Detection
The system must detect when performance degrades and requires policy updates.
Rationale:
Detecting performance degradation ensures timely interventions to maintain safety.
Primary indicator of safety is an increase in Constraint Violation Rate (CVR)
Detection of reward trend deviations such as continuous downward trends or rolling average drops beyond a threshold over time
Safe fallback activation is triggered when degradation is detected.
P3
Transparency
REQ 15
Post-Training Performance Audits
The RL system must facilitate structured audits post-training to evaluate decision-making quality and fairness.
Rationale:
Ensuring transparency in model updates and fairness evaluations strengthens accountability and trust in AI-driven decision-making.
Implement mechanisms for:
• Model confidence assessments
• Adversarial scenario testing
• Bias detection
Detection of reward trend deviations such as continuous downward trends or rolling average drops beyond a threshold over time
Safe fallback activation is triggered when degradation is detected.
Trusthworthy RL
UC2
P4
Robustness
REQ 1
Prioritization of High-Risk Scenarios
The system shall prioritize high-risk or low-probability scenarios in testing and replay strategies.
Rationale:
Ensuring that the system frequently encounters high-risk scenarios enhances its robustness to critical and rare situations.
Identify high-risk situations with input from human experts and prioritize them based on their impact.
Use a function quantifying scenario impact to construct a probability distribution that prioritizes high-risk scenarios.
P5
Privacy
REQ 25
Sanitization and Anonymization of Sensitive Data System
During data preprocessing, the system shall strip GPS and other identifying metadata to ensure anonymization.
Rationale:
Removing identifying metadata reduces the risk of data exposure and ensures privacy compliance.
Remove GPS coordinates and any identifying metadata from training data before storage and processing.
Ensure no personally identifiable information (PII) is included in datasets.
P6
AI Expertise
REQ 11
Intuitive Information Presentation for Human Oversight
The system shall present real-time information in a manner that aligns with the expertise and cognitive load of the overseer.
Rationale:
Ensuring information is clearly presented improves decision-making and reduces cognitive overload for the overseer.
Timeline visualization of past actions and predicted trajectory.
Diagram of the robot model with highlighted visual information for CoG and foot contacts.
Real-time dashboard with visual indicators for key state variables such as velocity, current, and policy confidence score.
Simplified color-coded alerts:
• Green for normal
• Yellow for caution
• Red for critical
Configurable verbosity levels:
• High-detail logs for engineers
• Summaries for operators

Fig. 5: We select one requirement for each of the six participants, where each was generated from that participants' responses to the ASKs associated an objective. We show the traceability for each requirement to related objectives and design decisions.

TABLE IV: Shows the requirement counts for the participant's original set (baseline), the generated ones (new and modified), and the participant's reviews of their 4 randomly selected ones.

	P1	P2	P3	P4	P5	P6
Original	7	11	10	10	21	10
Generated	18	20	21	14	8	12
New	14	19	16	14	4	10
Modified	4	1	5	0	4	2
Review						
Accept	4	3		1	4	2
Minor Mod.		1	4	1		2
Major Mod.				1		
Rejected				1		

that at least four new requirements were added in all cases, with some participants receiving as many as 19 additions. The smallest increase was observed for P5, likely due to their original requirements containing nearly twice as many as other participants, suggesting the framework's impact may vary based on a user's starting point.

In addition to the quantitative results, two of the researchers independently examined interview feedback using a bottom-up approach to identify key themes, allowing insights to emerge organically rather than being imposed by predefined categories. They discussed the themes and agreed on how to best organize the feedback. As a result, several key themes emerged: (a) the framework surfaced issues users hadn't previously considered, (b) revealed new details about known issues, (c) prompted requests for additional ASKs, (d) highlighted specific RL aspects that stood out to participants, and (e) exposed pain points in the process

Additional Considerations: Participants reported that the framework helped surface issues they had not previously considered. For example, P1 stated that "I think my requirements captured around 30% of the questions in [Task 1]. 70% of the questions I hadn't thought about." They elaborated, saying "Some of the topics I was less familiar with, and I did not know in as much depth. I had to research on my own to figure out answers to the questions.".

Missing Details: In a second closely related theme, several participants acknowledged that, while some topics were present in their initial requirements, they lacked sufficient detail. Many of the participants expressed that the ASKs required more critical thinking and elicited greater detail than their original requirements. For example, P2 noted that they "would have provided more details" in Task 1 "without assuming that some knowledge was standard."

Suggested ASKs: While the ASKs helped participants introduce new considerations into their requirements, four out of six felt that they needed further expansion to be truly comprehensive. P1 recommended two additions focused on incorporating external data sources and enhancing robustness through sensor redundancy. P3 felt that the topic of "fairness in signal distribution" was not adequately elicited; while P4 emphasized the need for a greater focus on "the learning process itself, such as which algorithm to use, how to address learning challenges like overestimation and underestimation in reinforcement learning algorithms, or how to design a good reward function." Finally, P5 suggested incorporating hardware-related considerations, such as sensor calibration and hardware degradation, as well as energy management concerns.

Highlighted RL aspects: Several participants emphasized that the framework was particularly helpful in highlighting key RL considerations related to safety, real-world conditions, and human oversight. In other words, it helped them consider how the RL agents would operate safely within the CPS. For example, P5 noted that the framework "required integrating RL's exploratory nature with strict, real-time safety constraints." Similarly, P3 appreciated the emphasis on safety, especially in addressing "the challenge of determining when the system should intervene." For real-world conditions, P3 highlighted that the framework "demanded a thorough analysis of how the system would operate in real-world scenarios, considering edge cases and bias detection." P1 also found this focus thought-provoking, particularly in relation to "formal verification systems, incorporating real-world feedback, and computational complexity."

Finally, participants appreciated the emphasis on human oversight. P6 noted that "the importance of human oversight tools and interpretability features became more apparent," while P3 acknowledged that it "highlighted the need for human operators to intervene when necessary."

Pain points: Two pain points in the *QuestRL Framework* were identified in our analysis, specifically redundancy and ambiguity in some ASKs.

"I felt that some questions were redundant," said P1, referring to three robustness-related ASKs. "All three have basically the same answer. I grouped them into one." While the ASKs were designed to be specific and address issues from multiple angles, this feedback highlights a downside. A simple solution could be to allow practitioners to link ASKs to previous responses, reducing redundancy and preventing unnecessary repetition when an ASK has already been addressed elsewhere (see Section VII).

Meanwhile, P4 found some ASKs too vague. For instance, they questioned ASK 2 ("How is the learning architecture optimized for the available computational resources?"), asking whether this referred to RL frameworks or neural networks. They also suggested rewording ASKs for clarity, such as revising "Have differences between the simulated and real environments been identified and addressed?" to "How can the gap between the simulated and real environment be identified and addressed?"

Conclusions: In response to RQ1, both quantitative and qualitative data indicate that *QuestRL* significantly helped participants consider a much broader range of design decisions, leading to more detailed specifications. The extent of this benefit—whether by introducing entirely new considerations or by prompting deeper critical thinking—varied depending on each participant's initial baseline.

At the same time, despite helping participants to address gaps in their own requirements, some noted that the ASKs themselves had gaps, presenting opportunities for refinement to improve coverage which we discuss in Section VII. Additionally, the evaluation highlighted that enhancing the clarity of certain ASKs could further improve their effectiveness in requirements elicitation.

B. RQ2: To what extent are requirements generated by QuestRL acceptable to RL-Engineers?

To answer this question, we examined the responses provided in Task 3, where participants rated four requirements generated using their design decisions. As shown in the lower half of Table IV, two participants marked all four requirements as acceptable, three others marked all of them as either acceptable or needing only minor modifications, and one (P4) graded them across the full spectrum from Acceptable to Rejected.

One reason for the requested changes was to reword the requirements to better align with conventional requirement phrasing. For instance, one generated requirement stated, "The system shall incorporate methods to handle sensor noise..." However, the participant suggested revising it to, "When operating, the robot shall perform well in the presence of sensor and actuator noise." Another reason for requesting modifications was that the LLM had updated an existing requirement by incorporating a design consideration from the ASKs but had omitted an assumption the participant deemed essential for proper interpretation. The remaining modification requests were suggestions to make the requirements more specific by filling in a missing detail, such as defining "RL policy confidence score". In the single case of outright rejection, the participant determined that the requirement was more appropriately categorized as an "operational constraint" rather than a requirement in itself.

Based on these results we are able to answer RQ2 by stating that the majority of the sampled requirements (22/24 or 92%) were found to be acceptable. In all instances, participants found the information provided value but felt that adjustments were necessary for optimal integration into their requirements.

C. RQ3: How do developers perceive the QuestRL Framework compared to manually creating requirements from scratch?

To explore this question, we analyzed responses to question #5, which compared Task 1 (manual requirements) and Task 2 (using *QuestRL*). Five out of six participants found Task 2 to be more challenging. Each of them attributed this to the need for deeper thinking. For example, P1 explained, "I needed to make sure that I could produce detailed answers and so there was an emphasis on making sure I understand the underlying problem and the potential solution for it." P2 echoed this, saying that "[Task 2] is more detailed than the design presented in Task 1," while P6 noted that it "demanded intricate technical trade-offs."

In contrast, one participant found Task 1 more difficult, explaining that "it was highly open-ended, requiring the formulation of system requirements from scratch. Unlike [Task 2], which provided specific questions to guide analysis."

Therefore, we conclude that participants perceive the *QuestRL Framework* as challenging the depth of thought and level of detail required compared to crafting their own requirements. Additionally, in one case, *QuestRL* was viewed as providing more structure to guide the analysis.

V. Related-Work

With the increasing adoption of AI-driven critical systems, ensuring their safety has become a key concern. To this end, both ISO/IEC 23894:2023 [9] and the NIST AI Risk Management Framework (AI RMF) [10] provide guidelines for managing AI risks, while researchers have also proposed various actionable practices for developing trustworthy AI [41]–[43]. Furthermore, safety assurance cases have been explored as a means of supporting AI safety, with frameworks such as Rueß and Burton's structured safety argument approach, which incorporates goals, assumptions, and solutions [44]. While these efforts primarily focus on establishing high-level, generalizable best practices, our *QuestRL Framework* is designed to complement these approaches by bridging the gap between abstract principles and concrete technical considerations specific to RL.

Researchers have also made progress within requirements engineering for AI by identifying six key areas for developing human-centered AI requirements [45]. While our work aligns with these areas, we extend beyond human-AI interaction to address RL-specific concerns, such as environment alignment.

In addition, several checklist-based approaches provide structured frameworks to address safety concerns in AI systems. Examples include the AI Safety and Security Checklist by HackerOne [46], which focuses on preventing harmful content and ensuring system robustness, and the LivePerson AI System Safety Checklist [47], which addresses risks like bias and security. Although these checklists are designed to be broadly applicable across different ML techniques, more targeted efforts have emerged, such as a recent RL-specific safety checklist [22]. While we draw on this work to help identify relevant ASKs, *QuestRL* moves beyond static checklists by adopting a question-driven approach that encourages critical thinking, enables traceability to system objectives, and facilitates downstream risk analysis and safety verification.

Within the RL domain, prior research has sought to create more trustworthy RL by mapping various methodologies to key objectives such as safety, robustness, and generalizability [21]. Our work builds on this foundation by incorporating additional objectives derived from AI risk management frameworks and structuring them into a process that supports system design and requirements elicitation.

Finally, we draw upon prior studies that have identified key challenges in RL development [19], [23]–[26], [48] and verification [8], [20]. We leverage these insights to design a set of ASKs that guide practitioners in addressing these challenges proactively, ensuring that critical considerations are incorporated from the start.

VI. Threats to Validity

Internal validity refers to the rigor of the experimental design and the extent to which results are attributable to interventions rather than confounding factors. Our small sample of six experts, all sourced from Upwork, may limit perspective and demographic diversity while introducing confirmation bias risk. We mitigate this by selecting experts with diverse backgrounds within RL, though future work should conduct additional experiments with more participants. Additionally, only a few LLM-generated requirements were validated by experts. To mitigate this, we randomly selected them across all different objectives to maximize coverage. Finally, our lightweight literature review approach may have missed relevant work due to its ad-hoc nature. However, we designed the framework to be extendable to additional RL considerations that future work may uncover in a more robust literature review. *External validity* concerns the generalizability of our findings. We evaluated only two different systems, which limits our ability to generalize findings across diverse systems. To address this, we deliberately selected two systems with distinct technical characteristics based on real-world cases from existing research. Furthermore, our method relies on only one LLM (GPT-4), which affects reproducibility and result interpretation. Our choice of GPT-4 reflects its widespread adoption and superior performance on similar tasks, though future work should evaluate additional models. *Construct validity* reflects how well the study measures intended variables and concepts. A potential threat is participants' use of LLMs during assigned tasks, as prior research shows widespread ChatGPT usage on crowdsourced platforms [49]. We assessed this risk by questioning participants about LLM usage—all but one indicated *occasional* use. Given ChatGPT's widespread adoption as a support tool, this likely reflects realistic baseline conditions rather than result distortion.

VII. Conclusions and Future Work

In this paper, we introduce the *QuestRL Framework*, a structured, question-driven approach designed to enhance the safety and reliability of RL in CPS. By breaking down high-level AI safety objectives into concrete RL-specific questions, the *QuestRL Framework* provides practitioners with a systematic method for defining and refining safety requirements early in the development process. Through two case studies in distinct CPS scenarios, where six RL experts applied the framework, we demonstrate how the *QuestRL Framework* fosters critical thinking, improves requirements coverage, and facilitates a more structured approach to safety assurance. Our findings highlight the framework's effectiveness in guiding practitioners toward more comprehensive and proactive requirement formulation, ultimately supporting the development of safer and more trustworthy RL-based systems. At the same time, they reveal opportunities for refinement and expansion. Based on participant feedback and observed limitations, we identify several promising directions for future work:

1. **Expand RL Considerations**

   A common recommendation from participants was to extend the framework to cover additional ASKs. Based on their feedback, we propose incorporating six additional ASKs in future iterations of the framework to provide more comprehensive coverage of the objectives. In Table V, we introduce each ASK beneath the primary objective that it would support.

TABLE V: Displays the 6 new questions that we propose based on our evaluation.

	Maintainability
1.	How does the RL agent balance short-term energy efficiency with long-term system sustainability?
2.	How does the RL system adapt to sensor drift and degradation over time?
	Accuracy
3.	What external data sources have been incorporated to enhance decision-making and how is it ensured that such data is reliable, up-to-date, and contextually relevant?
4.	What strategies are used to mitigate estimation biases (i.e. overestimation and underestimation) in the agent's decision making?
5.	How does the selection of RL architecture align with the assumptions of the system such as data availability and learning objectives?
	Robustness
6.	What mechanisms are in place to detect and mitigate inconsistencies or sensor failures?

2. **Adapt Framework to Other AI/ML Paradigms** While *QuestRL* is specifically designed for RL challenges, the underlying question-driven methodology shows potential for adaptation to other AI/ML paradigms. Given that the framework already incorporates core AI trustworthiness objectives, future research could explore the development of new ASKs that target the distinct safety challenges posed by alternative machine learning approaches.
3. **Enhance Tool Support**

   Traceability analysis can be significantly enhanced through the use of graphical visualizations and specialized tools [16], [50]. To further improve the process, we propose incorporating a requirements management or traceability tool into the framework in future iterations. In addition to improving requirements and trace link analysis, such a tool could address concerns about redundant responses by allowing practitioners to easily link responses when they address multiple ASKs.

While the *QuestRL Framework* already supports RL practitioners in addressing safety requirements, the future directions we outline aim to extend its utility—advancing efforts to operationalize safety and trust in a wider range of AI-based systems. To support continued research and replication, we provide supplementary materials, including evaluation data and framework artifacts, at [1].

References

[1] K. R. Dearstyne, P. A. Granadeno, T. Chambers, and J. Cleland-Huang, "Questrl: Supplementary materials," https://zenodo.org/records/15825459, 2025, accessed: 2025-07-06.

[2] R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction, 2nd ed. Cambridge, MA: MIT press, 2018.

[3] S. Aradi, "Survey of deep reinforcement learning for motion planning of autonomous vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 2, pp. 740–759, 2020.

[4] D. Zhang, X. Han, and C. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids," CSEE Journal of Power and Energy Systems, vol. 4, no. 3, pp. 362–370, 2018.

[5] X. Liu, H. Xu, W. Liao, and W. Yu, "Reinforcement learning for cyber-physical systems," in 2019 IEEE International Conference on Industrial Internet (ICII), 2019, pp. 318–327.

[6] T. Rupprecht and Y. Wang, "A survey for deep reinforcement learning in markovian cyber–physical systems: Common problems and solutions," Neural Networks, vol. 153, pp. 13–36, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608022001873

[7] H. Mirzaei Buini, S. Peter, and T. Givargis, "Adaptive embedded control of cyber-physical systems using reinforcement learning," IET Cyber-Physical Systems: Theory & Applications, vol. 2, 07 2017.

[8] P. van Wesel, "Challenges in the Verification of Reinforcement Learning Algorithms," 2017.

[9] "ISO/IEC 23894:2023." [Online]. Available: https://www.iso.org/standard/77304.html

[10] E. Tabassi, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," NIST, Jan. 2023, last Modified: 2023-01-26T08:01-05:00 Publisher: Elham Tabassi. [Online]. Available: https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10

[11] R. Salehin, "Missing Requirements Information and its Impact on Software Architectures: A Case Study," 2013. [Online]. Available: https://www.semanticscholar.org/paper/Missing-Requirements-Information-and-its-Impact-on-Salehin/bbf448ab8a7684f1c8bc944ad9b8c92fdb85b20a

[12] M. Kauppinen, M. Vartiainen, J. Kontio, S. Kujala, and R. Sulonen, "Implementing requirements engineering processes throughout organizations: success factors and challenges," Information and Software Technology, vol. 46, no. 14, pp. 937–953, Nov. 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950584904000692

[13] J. J. Y. Tan, K. N. Otto, and K. L. Wood, "Relative impact of early versus late design decisions in systems development," Design Science, vol. 3, p. e12, Jan. 2017. [Online]. Available: https://www.cambridge.org/core/journals/design-science/article/relativeimpact-of-early-versus-late-design-decisions-in-systemsdevelopment/10B13C6901E51E7F6337250A9CA36E17

[14] J. Cleland-Huang, O. Gotel, J. H. Hayes, P. Mäder, and A. Zisman, "Software traceability: trends and future directions," in Proceedings of the on Future of Software Engineering, FOSE 2014, Hyderabad, India, May 31 - June 7, 2014, J. D. Herbsleb and M. B. Dwyer, Eds. ACM, 2014, pp. 55–69. [Online]. Available: https://doi.org/10.1145/2593882.2593891

[15] K. R. Dearstyne, A. D. Rodriguez, and J. Cleland-Huang, "Supporting Software Maintenance with Dynamically Generated Document Hierarchies," Aug. 2024, arXiv:2408.05829 [cs]. [Online]. Available: http://arxiv.org/abs/2408.05829

[16] ——, "ROOT: Requirements Organization and Optimization Tool," in 2024 IEEE International Conference on Software Maintenance and Evolution (ICSME), Oct. 2024, pp. 883–887, iSSN: 2576-3148. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10795032

[17] K. Johnson and N. G. Leveson, "Investigating safety and cybersecurity design tradespace for manned-unmanned aerial systems integration using systems theoretic process analysis," in 44. Jahrestagung der Gesellschaft für Informatik, Big Data - Komplexität meistern, INFORMATIK 2014, Stuttgart, Germany, September 22-26, 2014, ser. LNI, E. Plödereder, L. Grunske, E. Schneider, and D. Ull, Eds., vol. P-232. GI, 2014, pp. 643–647. [Online]. Available: https://dl.gi.de/handle/20.500.12116/2960

[18] G. Dulac-Arnold, N. Levine, D. J. Mankowitz, J. Li, C. Paduraru, S. Gowal, and T. Hester, "Challenges of real-world reinforcement learning: definitions, benchmarks and analysis," Machine Learning, vol. 110, no. 9, pp. 2419–2468, Sep 2021. [Online]. Available: https://doi.org/10.1007/s10994-021-05961-4

[19] P. Hamadanian, M. Schwarzkopf, S. Sen, and M. Alizadeh, "How Reinforcement Learning Systems Fail and What to do About It," 2022.

[20] S. Sharma, M. A. B. U. Rahim, S. Hussain, M. R. Abid, and T. Liu, "Using Deep Reinforcement Learning And Formal Verification in Safety Critical Systems: Strategies and Challenges," in 2023 IEEE 23rd International Conference on Software Quality, Reliability, and Security Companion (QRS-C), Oct. 2023, pp. 834–842, iSSN: 2693-9371. [Online]. Available: https://ieeexplore.ieee.org/document/10429991/

[21] M. Xu, Z. Liu, P. Huang, W. Ding, Z. Cen, B. Li, and D. Zhao, "Trustworthy Reinforcement Learning Against Intrinsic Vulnerabilities: Robustness, Safety, and Generalizability," Sep. 2022, arXiv:2209.08025. [Online]. Available: http://arxiv.org/abs/2209.08025

[22] T. Yamagata and R. Santos-Rodriguez, "Safe and Robust Reinforcement Learning: Principles and Practice," Mar. 2024, arXiv:2403.18539 [cs] version: 2. [Online]. Available: http://arxiv.org/abs/2403.18539

[23] J. Ibarz, J. Tan, C. Finn, M. Kalakrishnan, P. Pastor, and S. Levine, "How to train your robot with deep reinforcement learning: lessons we have learned," The International Journal of Robotics Research, vol. 40, no. 4-5, pp. 698–721, Apr. 2021, publisher: SAGE Publications Ltd STM. [Online]. Available: https://doi.org/10.1177/0278364920987859

[24] M. M. Morovati, F. Tambon, M. Taraghi, A. Nikanjam, and F. Khomh, "Common challenges of deep reinforcement learning applications development: an empirical study," Empir Software Eng, vol. 29, no. 4, p. 95, Jun. 2024. [Online]. Available: https://doi.org/10.1007/s10664-024-10500-5

[25] S. Gu, L. Yang, Y. Du, G. Chen, F. Walter, J. Wang, and A. Knoll, "A Review of Safe Reinforcement Learning: Methods, Theories, and Applications," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 12, pp. 11 216–11 235, Dec. 2024, conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10675394

[26] Z. Zhou, G. Liu, and Y. Tang, "Multi-Agent Reinforcement Learning: Methods, Applications, Visionary Prospects, and Challenges," IEEE Trans. Intell. Veh., pp. 1–23, 2024, arXiv:2305.10091 [cs]. [Online]. Available: http://arxiv.org/abs/2305.10091

[27] K. Dearstyne, P. T. A. Granadeno, T. Chambers, and J. Cleland-Huang, "Poster: Evaluating reinforcement learning safety and trustworthiness in cyber-physical systems," in Proceedings of the 4th International Conference on AI Engineering – Software Engineering for AI (CAIN), Co-located with ICSE 2025. ACM, April 2025, poster Presentation. [Online]. Available: https://arxiv.org/abs/2503.09388

[28] T. T. Tun, R. R. Lutz, B. Nakayama, Y. Yu, D. Mathur, and B. Nuseibeh, "The role of environmental assumptions in failures of DNA nanosystems," in 1st IEEE/ACM International Workshop on Complex Faults and Failures in Large Software Systems, COUFLESS 2015, Florence, Italy, May 23, 2015, 2015, pp. 27–33. [Online]. Available: http://dx.doi.org/10.1109/COUFLESS.2015.12

[29] M. Rahimi, W. Xiong, J. Cleland-Huang, and R. R. Lutz, "Diagnosing assumption problems in safety-critical products," in Proceedings of the 32nd IEEE/ACM International Conference on Automated Software Engineering, ASE 2017, Urbana, IL, USA, October 30 - November 03, 2017, G. Rosu, M. D. Penta, and T. N. Nguyen, Eds. IEEE Computer Society, 2017, pp. 473–484. [Online]. Available: https://doi.org/10.1109/ASE.2017.8115659

[30] A. van Lamsweerde, Requirements Engineering: From System Goals to UML Models to Software Specifications. Wiley, 2009.

[31] R. Lutz, A. Patterson-Hine, S. Nelson, C. R. Frost, D. Tal, and R. Harris, "Using obstacle analysis to identify contingency requirements on an unpiloted aerial vehicle," Requirements Engineering, vol. 12, no. 1, pp. 41–54, 2007.

[32] DOT/FAA/AR-08/32, Requirements Engineering Management Handbook, 2009.

[33] B. Scientific, "Pacemaker system specification," http://sqrl.mcmaster.ca/_SQRLDocuments/PACEMAKER.pdf, Boston Scientific, 2007.

[34] J. P. L. U. S. R. Board and J. Casani, Report on the loss of the Mars polar lander and Deep Space 2 missions. Jet Propulsion Laboratory, California Institute of Technology, 2000.

[35] N. Marques, R. R. Silva, and J. Bernardino, "Using ChatGPT in Software Requirements Engineering: A Comprehensive Review," Future Internet, vol. 16, no. 6, p. 180, Jun. 2024, number: 6 Publisher:

Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/1999-5903/16/6/180

[36] L. Bencheikh and N. Höglund, "Exploring the efficacy of chatgpt in generating requirements: An experimental study," Aug. 2023. [Online]. Available: https://gupea.ub.gu.se/handle/2077/77957

[37] G. Zhang, F. Chang, J. Jin, F. Yang, and H. Huang, "Multi-objective deep reinforcement learning approach for adaptive traffic signal control system with concurrent optimization of safety, efficiency, and decarbonization at intersections," Accident Analysis & Prevention, vol. 199, p. 107451, May 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457523004980

[38] L. Smith, I. Kostrikov, and S. Levine, "A Walk in the Park: Learning to Walk in 20 Minutes With Model-Free Reinforcement Learning," Aug. 2022, arXiv:2208.07860 [cs]. [Online]. Available: http://arxiv.org/abs/2208.07860

[39] A. Mavin, P. Wilkinson, A. Harwood, and M. Novak, "Easy Approach to Requirements Syntax (EARS)," in Proceedings of the 2009 17th IEEE International Requirements Engineering Conference, RE, ser. RE '09. USA: IEEE Computer Society, Aug. 2009, pp. 317–322. [Online]. Available: https://doi.org/10.1109/RE.2009.9

[40] OpenAI, "GPT-4 Technical Report," Mar. 2023, arXiv:2303.08774 [cs]. [Online]. Available: http://arxiv.org/abs/2303.08774

[41] A. Serban, K. van der Blom, H. Hoos, and J. Visser, "Practices for engineering trustworthy machine learning applications," in 2021 IEEE/ACM 1st Workshop on AI Engineering - Software Engineering for AI (WAIN). IEEE Press, 2021, p. 97–100. [Online]. Available: https://doi.org/10.1109/WAIN52551.2021.00021

[42] Q. Lu, L. Zhu, X. Xu, J. Whittle, and Z. Xing, "Towards a roadmap on software engineering for responsible ai," in Proceedings of the 1st International Conference on AI Engineering: Software Engineering for AI, ser. CAIN '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 101–112. [Online]. Available: https://doi.org/10.1145/3522664.3528607

[43] C. S. Wickramasinghe, D. L. Marino, J. Grandio, and M. Manic, "Trustworthy AI Development Guidelines for Human System Interaction," in 2020 13th International Conference on Human System Interaction (HSI), Jun. 2020, pp. 130–136, iSSN: 2158-2254. [Online]. Available: https://ieeexplore.ieee.org/document/9142644

[44] Fraunhofer Institute for Cognitive Systems (IKS), "Safe ai: How is this possible?" 2021, accessed: 2024-11-16. [Online]. Available: https://www.iks.fraunhofer.de/content/dam/iks/documents/whitepaper-safeai.pdf

[45] K. Ahmad, M. Abdelrazek, C. Arora, A. Agrahari Baniya, M. Bano, and J. Grundy, "Requirements engineering framework for human-centered artificial intelligence software systems," Applied Soft Computing, vol. 143, p. 110455, Aug. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494623004738

[46] HackerOne, "Ai safety and security checklist," 2024, accessed: 2024-11-16. [Online]. Available: https://www.hackerone.com/ai-security-checklist

[47] LivePerson, "Ai system safety checklist," 2024. [Online]. Available: https://www.liveperson.com/resources/reports/ai-safety-checklist/,note={Accessed:2024-11-16}

[48] G. Dulac-Arnold, N. Levine, D. J. Mankowitz, J. Li, C. Paduraru, S. Gowal, and T. Hester, "Challenges of real-world reinforcement learning: definitions, benchmarks and analysis," Mach Learn, vol. 110, no. 9, pp. 2419–2468, Sep. 2021. [Online]. Available: https://doi.org/10.1007/s10994-021-05961-4

[49] V. Veselovsky, M. H. Ribeiro, and R. West, "Artificial artificial artificial intelligence: Crowd workers widely use large language models for text production tasks," 2023. [Online]. Available: https://arxiv.org/abs/2306.07899

[50] J. Cleland-Huang and R. Habrat, "Visual Support In Automated Tracing," in Second International Workshop on Requirements Engineering Visualization (REV 2007), Oct. 2007, pp. 4–4. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4473004