Id	Question	Response	Example
A1	Scope		
A1.1	What are the primary **objectives** that the RL agent should accomplish?		Maximize throughput and minimize latency in an industrial robotic assembly line by dynamically adjusting movement patterns based on sensor feedback.
A2	Users		
A2.1	What is the level of **AI expertise** of the user(s)?	Rate Expertise	**Beginner**: The user has little to no experience with AI, possibly unfamiliar with basic concepts or terminology. **Novice**: The user has a basic understanding of AI, but may still need guidance and support in applying AI techniques to real-world problems. **Intermediate**: The user has a solid understanding of AI and can implement standard models, but may require help with more complex or specialized tasks. **Advanced**: The user is skilled in AI, capable of developing and optimizing complex models, and can handle most AI-related challenges independently. **Expert**: The user has deep, specialized knowledge in AI, often contributing to research or development of new methodologies and technologies.
A2.2	What is the level of **interaction** of the user(s) with the agent?	Rate Level of Interaction	**Passive**: The user is not actively involved with the agent, merely observing its operation or outcomes without direct interaction. **Guided**: The user oversees the agent's deployment, monitoring its performance and providing feedback or adjustments during its operation. **Contributor**: The user is actively involved in the creation or improvement of the agent, such as training the model or building its system components.
A3	Learning		
A3.2	**Online** (adaption during interactions with environment) or **Offline** (learns solely from pre-collected dataset)?	Select	**Online**: A self-driving car continuously adapts to new road conditions, such as unexpected construction zones or weather changes, by updating its policy in real-time. **Offline**: A predictive maintenance system for industrial machinery is trained on a historical dataset of sensor readings and failure events, making inferences without further updates during deployment
A3.3	Task Persistence: **Persistent** (agent can retry the task indefinitely) or **non-persistent with reset** (task can be retried, but only after a significant reset of the environment) or **irreversible** (agent can only perform task once)?	Select	**Persistent**: A robotic arm in a manufacturing line repeatedly attempts to grasp and assemble components, refining its strategy over hundreds of thousands of trials. **Non-persistent with reset**: A robotic barista preparing a drink must clean and reset the station before making another cup. **Irreversible**: A robotic chef slicing vegetables permanently alters their shape, preventing retries on the same item.
A3.1	Reward		
A3.1.1	**Immediate** (feedback is provided right after an action is taken) or **Delayed** (feedback may span multiple actions or time steps)?	Select	**Immediate**: A robotic gripper receives a reward signal immediately after successfully picking up an object. **Delayed**: A delivery drone receives feedback only after it has successfully completed its entire delivery route, including takeoff, navigation, and drop-off, even though there were many intermediate actions involved. A
A3.1.2	**Stochastic** (feedback contains randomness or uncertainty for a given action) or **Deterministic** (feedback is consistent and deterministic for a given action)?	Select	**Stochastic**: A robot navigating a cluttered environment receives varying feedback on its movement, depending on factors like object interaction or sensor noise. **Deterministic**: A robotic arm consistently receives the same feedback after executing a precise action, such as placing an object in a fixed position.
A3.1.3	**Sparse** (feedback is only given occasionally or in specific situations) or **Dense** (feedback is frequent and consistently available)?	Select	**Sparse**: A robot playing chess receives feedback only when it wins or loses a game, not after each individual move. **Dense**: A robot playing chess receives feedback after every move it makes, adjusting its strategy constantly based on the outcome of each individual action.

A4	**Environment**			
A4.1	**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)?	Select	**Fully observable:** A robotic arm in a controlled factory environment has access to complete information about its surroundings, such as the exact position of all components and tools. **Partially observable**: A self-driving car has limited visibility due to obstacles or weather conditions, so it must rely on sensors and predictions to fill in gaps in its understanding of the environment.	
A4.2	State Transitions are **Deterministic** (completely predictable given the current state and action) or **Stochastic** (involving randomness or uncertainty)?	Select	**Deterministic**: A robotic arm moving along a fixed path in a factory will always reach the same destination given the same starting point and action, with no randomness involved. **Stochastic**: A drone navigating through varying wind conditions experiences unpredictable changes in its flight path, even if the action (e.g., moving forward) is the same.	
A4.3	State Space is: **Discrete** (finite set of states) or **Continuous** (infinite possible states) or **High-dimensional** (many variables contributing to the state)?	Select	**Discrete**: A grid-based robot navigation system, where the robot can only occupy specific grid points, such as a robot moving on a chessboard. **Continuous**: A self-driving car's position and velocity, where the car can be in any position and have any speed within a continuous range. **High-dimensional**: A drone navigating a visual scene, where the state includes multiple objects' positions and the drone's own location and orientation.	
A4.4	Training + testing distribution is: **IID** (drawn from the same distribution) or **OOD** (drawn from different distributions) or **Non-stationary** (drawn from time-variant, non-stationary distributions)?	Select	**IID:** A robot trained and tested on data from the same environment, such as a robotic arm in a factory using consistent conditions for both training and testing. **OOD**: A self-driving car trained on data from urban areas and tested in rural environments, where the distributions of road types and conditions differ. **Non-stationary**: A robot trained to navigate in a warehouse with seasonal changes in inventory, causing the environment to evolve over time.	
A4.5	Action space is: **Discrete** (a finite set of possible actions) or **Continuous** (an infinite range of possible actions)?	Select	**Discretee**: A robot can only perform a set of predefined actions, such as "pick up," "move forward," or "rotate 90 degrees," with a finite number of choices. **Continuous**: A robotic arm can adjust its position at any angle or with any speed, allowing for an infinite range of possible actions.	
A4.6	**Single-agent** or **Multi-agent (cooperation vs. competition)**?	Select	**Single-agent**: A robotic vacuum operates independently, navigating a home and performing cleaning tasks without interacting with other agents. **Multi-agent (cooperation):** Multiple drones work together to deliver packages, coordinating their paths to avoid collisions and optimize delivery time. **Multi-agent (competition):** In a competitive robot soccer game, each robot tries to score goals while preventing the opposing team from doing so.	
A4.7	What physical constraints must the agent adhere to during operation? These might include temperature limits, workspace boundaries, max-velovity, etc.		The agent must operate within a temperature range of -10°C to 50°C, stay within a 10-meter workspace, and not exceed a maximum velocity of 5 m/s.	
A5	**Hardware/Platform**			
A5.1	Will the agent be executed **locally** on the device, or will it be run **remotely** on a server or cloud?	Select	**Locally on the device:** An autonomous vacuum cleaner processes sensor data and makes navigation decisions directly on-board without needing a connection to a remote server. **Remotely on a server:** A self-driving car may rely on a cloud server for real-time map updates and traffic data, while using onboard sensors for immediate navigation decisions.	
A5.2	What is the available **processing power** for the RL agent on the device?		The RL agent has 1.2 GHz processing power on a robotic vacuum with an embedded ARM processor.	
A5.3	What is the total available **RAM** and **storage** on the device for running the RL agent?		The RL agent has 4 GB of RAM and 64 GB of storage on a robotic vacuum with an embedded system.	

A5.4	What is the maximum acceptable **latency** (in milliseconds) for the RL agent to process inputs and produce outputs?		The RL agent must process inputs and produce outputs within 100 milliseconds on a robotic vacuum to ensure real-time navigation and obstacle avoidance.	
A5.5	Will the RL agent be deployed across **multiple** devices with **varying hardware configurations**, or will it run on a **uniform** set of hardware? (Note: if using multiple, assumptions should be provided for each different hardware configuration)	Select	**Uniform set of hardware:** The RL agent runs on a fleet of delivery robots, all equipped with identical processors, sensors, and cameras. **Multiple devices with varying hardware configurations:** The RL agent operates across a range of devices, such as a mobile robot with limited processing power and a warehouse robot with a high-performance GPU, each tailored to their specific capabilities.	