

COURSE GLOSSARY

Introduction to AI Agents

- **AI Agent:** An autonomous system that perceives its environment, makes decisions, and takes actions to achieve specific goals.
- **Autonomy:** The ability of an AI agent to operate and make decisions without continuous human intervention.
- **Reactive Agent:** An AI agent that responds directly to environmental stimuli without using memory or internal models.
- **Deliberative Agent:** An AI agent that uses reasoning, planning, and internal models to decide on actions.
- **Hybrid Agent:** An AI agent that combines reactive and deliberative approaches for more flexible behavior.
- **Perception:** The process of collecting information about the environment through sensors or data inputs.
- **Action:** The output or behavior performed by an AI agent to influence its environment.
- **Environment:** The external system or context in which an AI agent operates and interacts.
- **State:** A representation of the environment at a specific point in time.
- **Policy:** A set of rules or strategies that an AI agent follows to decide its actions.
- **Utility Function:** A mathematical function that measures how desirable a given state or outcome is for an agent.
- **Goal-based Agent:** An AI agent that selects actions to achieve predefined objectives.
- **Learning Agent:** An AI agent that improves performance over time by learning from experiences.
- **Multi-Agent System (MAS):** A system in which multiple AI agents interact, collaborate, or compete.
- **Communication (Agents):** The exchange of information between AI agents, often through predefined protocols.
- **Coordination:** The process of aligning actions between multiple agents to achieve collective goals.
- **Negotiation (Agents):** A decision-making process in which agents resolve conflicts and agree on outcomes.
- **Reinforcement Learning:** A machine learning approach where agents learn by receiving feedback in the form of rewards or penalties.
- **Exploration vs Exploitation:** The balance between trying new actions to gain knowledge (exploration) and using known actions to maximize rewards (exploitation).
- **Belief-Desire-Intention (BDI) Model:** A model of agent reasoning based on beliefs (knowledge), desires (objectives), and intentions (commitments).
- **Agent Architecture:** The structural design that defines how an AI agent's components (perception, reasoning, actions) are organized.
- **Agent Lifecycle:** The stages an AI agent goes through, including initialization, interaction, learning, and termination.
- **Trust in Agents:** The degree to which humans or other agents rely on an AI agent to act predictably and ethically.
- **Ethical AI Agent:** An agent designed to align its decisions and actions with human values, fairness, and safety guidelines.