

AI Agent Project Evaluation Grid

Part A: AI Agent Architecture & Implementation (10 points)

Skills	Observable	Points	Observable Criteria	Points Obtained
A01 - Design and implement multi-agent system with clear agent roles	Agent Architecture & Coordination	0	Single agent only or no clear agent implementation	
		2	Two agents exist but with unclear roles or poor coordination	
		3	Two distinct agents with defined roles but limited interaction	
		5	Multiple agents with clear roles and coordination	
A02 - Implement effective LLM integration with appropriate model selection	LLM Model Choice & Implementation	0	No LLM integration or hardcoded responses	
		2	Basic LLM integration	
		3	Appropriate model selection with basic implementation	
		5	Well-justified model choice with good prompt engineering (Jinja, PyDantic)	

Part B: Fairness & Bias Evaluation (15 points)

Skills	Observable	Points	Observable Criteria	Points Obtained
B01 - Identify and document potential	Bias Identification	0	No bias consideration mentioned	

biases in the system	& Documentation			
		2	Minimal bias discussion without specific identification	
		3	Basic identification of 1-2 potential biases	
		5	Comprehensive identification of 3+ biases with evidence	
B02 - Calculate and visualize fairness metrics	Fairness Metrics Implementation	0	No fairness metrics calculated	
		2	Attempted metrics but incorrect implementation	
		3	1-2 fairness metrics correctly calculated (e.g., demographic parity)	
		5	3+ metrics calculated with visualization	
B03 - Implement bias mitigation strategies	Mitigation Techniques	0	No mitigation attempted	
		2	Mitigation mentioned but not implemented	
		3	Basic mitigation attempt	
		5	Mitigation strategy fully implemented with results	

Bonus Points (up to 8 additional)

- Novel fairness metric or mitigation approach: +3 points
- Exceptional documentation or visualization: +2 points
- Open-source contribution from project: +3 points