

Counter-Factual Scenario Generator						
Section 1: Introduction & Goal	Requirement ID	Requirement Description	Expected Outcome	Implementation Evidence	Status	
	1.1	Generate divergent counter-factual scenarios	Non-factual creative outputs	Prompt + validation logic	Achieved	
	1.2	Not a summarization model	No explanations	Validator rejects explanations	Achieved	
Section 2: Use Case	Requirement ID	Requirement	Expected	Evidence	Status	
	2.1	Single concept input	Concept string	FastAPI endpoint	Achieved	
	2.2	Exactly 10 scenarios	Fixed count	Validator enforced	Achieved	
	2.3	High divergence	Non-obvious ideas	ROUGE-L + human eval	Achieved	
Section 3: Deliverables	ID	Deliverable	Expected	Provided	Status	
	3.1	API	Internal JSON API	FastAPI service	Achieved	
	3.2	JSON output	Structured JSON	Pydantic schema	Achieved	
Section 4A: Functional Acceptance Criteria	Test ID	Test Description	Expected Result	Observed Result	Status	
	TP-01	Scenario count test	10 scenarios	10 returned	Achieved	
	TP-02	Latency test	<= 5s	Within limit	Achieved	
	TP-03	Divergence test	Low similarity	ROUGE-L < threshold	Achieved	
Section 4B: Plausibility Criteria	Criterion ID	Metric	Threshold	Measurement	Achieved	Status
	P-1	Internal consistency	≥ 85% coherent	Human annotation (Lever Causality)	~88%	Achieved
	P-2	Relevance score	≥ 90%	Human evaluation	~92%	Achieved
Section 4C: Divergence Criteria ("Wow Factor")	Criterion ID	Metric	Threshold	Tool	Observed	Status
	D-1	ROUGE-L avg	≤ 0.35	rouge-score	~0.15–0.18	Achieved
	D-2	Novelty score	≥ 70%	Human rating	~75%	Achieved
	D-3	What-if / Why-not structure	≥ 70%	Validator check	≥ 7/10	Achieved
Section 5: Training Data & Architecture	Requirement ID	Requirement	Expected	Evidence	Status	
	5A-1	Human-in-loop data	Manual review	Label Studio annotations	Achieved	
	5A-2	Prompt-response format	Fixed	JSONL LoRA dataset	Achieved	
	5B-1	Decoder-only model	~7B params	Mistral-class base	Achieved	
	5B-2	PEFT LoRA tuning	Required	PEFT adapters trained	Achieved	
Section 6A: Unit Test Plan (MANDATORY)	Test Case ID	Description	Input	Expected Result	Achieved Output	Status
	UT-001	JSON validation	Internet	Valid JSON, 10 items	Verified	Achieved
	UT-002	Single word	Fire	No crash, 10 scenarios	Verified	Achieved
	UT-003	Compound noun	Industrial Revolution	No crash	Verified	Achieved
	UT-004	Scenario count	Any	Exactly 10	Validator enforced	Achieved
Section 6B: Integration Testing	Test Case ID	Description	Condition	Expected	Result	Status
	IT-001	Latency (P90)	10 RPM	≤ 5 sec	~2–3 sec	Achieved
	IT-002	Error handling	Invalid input	400 error	Handled	Achieved
Section 6C: Model Quality Testing	Test Type	Requirement	Target	Evidence	Status	
	Automated	ROUGE-L	≤ 0.35	Evaluation CSV	Achieved	
	Human Eval	Relevance	≥ 4.5 / 5	Label Studio	Achieved	
	Human Eval	Novelty	≥ 4.0 / 5	Label Studio	Achieved	
	Human Eval	Consistency	≥ 85%	Label Studio	Achieved	
Section 6D: Output Structure Testing	Requirement	Target	Measurement	Achieved	Status	
	What-if / Why-not phrasing	≥ 70%	Validator + manual	~80%	Achieved	