GenAI Solution Evaluation Rubric

Project: [Project Name]
Date: [Date]

Category	Criteria	1 - Poor	2 - Fair	3 - Good	4 - Excellent	Score
Model Quality	**Accuracy & Relevance**	Outputs are frequently inaccurate, irrelevant, or contain hallucinations. Requires constant human correction.	Outputs are generally on-topic but often lack precision or require significant editing.	Outputs are consistently accurate and relevant for common use cases. Minor edits may be needed.	Outputs are highly accurate, nuanced, and require minimal to no human correction, even on complex queries.	
Operational Safety	**Risk of Harmful Output**	The system could plausibly generate dangerously incorrect information (e.g., wrong safety procedures) without clear	The system provides general disclaimers but lacks specific guardrails for safety-critical topics.	The system identifies potentially sensitive queries and provides strong cautionary advice.	The system actively refuses to answer safety-critical queries or directs users to verified human-run procedures.	

		warnings.				
Perform ance	**Latency / Speed**	Responses take longer than 30 seconds, significantly disrupting user workflow.	Responses take 10-30 seconds, causing noticeable but manageable delays.	Responses are consistently delivered in 5-10 seconds, feeling responsive.	Responses are nearly instantaneous (<5 seconds), enabling a fluid, conversational experience.	
Cost	**Total Cost of Ownership (TCO)**	The solution is prohibitively expensive, with a projected ROI of less than 12 months.	The cost is significant, with a projected ROI of 18-24 months.	The cost is reasonable, with a clear and achievable ROI within 12-18 months.	The solution is highly cost-effective, with a projected ROI of under 12 months and low ongoing operational costs.	
Overall						_ / 16

Scoring Guide:

- **13-16 (Excellent):** Solution is ready for scaling. Strong business case and low risk.
- **9-12 (Good):** Solution is viable for production but may require minor improvements.
- **5-8 (Fair):** Significant concerns exist. Proceed with caution and address key weaknesses before wider deployment.
- **< 5 (Poor):** Solution is not ready. Re-evaluate the approach or technology.