(venv) C:\Main Files\Soorya\7thSem\LLM\_Prod\Exp2>python final\_eval.py

🚀 Initializing RAG Evaluator...

📊 Starting evaluation with queries from queries.json

🐌 Starting 3 RPM LIMITED evaluation of 20 queries...

Delay: 25 seconds between queries

Estimated time: 8.3 minutes

============================================================

[1/20] How can I disable SMS notification from dashboard?...

✅ Grade: A (Quality: 0.779, Latency: 3461ms)

⏳ Waiting 25s for VoyageAI rate limit...

[2/20] How long would it require to become P2M merchant a...

✅ Grade: A (Quality: 0.840, Latency: 2453ms)

⏳ Waiting 25s for VoyageAI rate limit...

[3/20] Why should you consider becoming a part of the Jio...

✅ Grade: A (Quality: 0.892, Latency: 2922ms)

⏳ Waiting 25s for VoyageAI rate limit...

[4/20] Why JioPay?...

✅ Grade: A (Quality: 0.738, Latency: 3168ms)

⏳ Waiting 25s for VoyageAI rate limit...

[5/20] Where is our Registered Address ?...

✅ Grade: B (Quality: 0.666, Latency: 2691ms)

⏳ Waiting 25s for VoyageAI rate limit...

[6/20] What is the Level 3 in the Complaint Resolution Ma...

✅ Grade: D (Quality: 0.296, Latency: 2602ms)

⏳ Waiting 25s for VoyageAI rate limit...

[7/20] What is the Remuneration to Non-Executive Director...

✅ Grade: A (Quality: 0.740, Latency: 2808ms)

⏳ Waiting 25s for VoyageAI rate limit...

[8/20] How is Intellectual Property Rights defined ?...

✅ Grade: B (Quality: 0.592, Latency: 2901ms)

⏳ Waiting 25s for VoyageAI rate limit...

[9/20] What is the advantage of using JioPay QR and who i...

✅ Grade: B (Quality: 0.619, Latency: 2530ms)

⏳ Waiting 25s for VoyageAI rate limit...

[10/20] What are JioPay Payment Gateway Products?...

✅ Grade: B (Quality: 0.590, Latency: 3133ms)

⏳ Waiting 25s for VoyageAI rate limit...

[11/20] How often the Grievance Redressal Policy Reviewed ...

✅ Grade: A (Quality: 0.775, Latency: 2500ms)

⏳ Waiting 25s for VoyageAI rate limit...

[12/20] What are the list of acceptable KYC Documents as a...

✅ Grade: A (Quality: 0.801, Latency: 2977ms)

⏳ Waiting 25s for VoyageAI rate limit...

[13/20] To what all data is the privacy policy applicable ...

✅ Grade: B (Quality: 0.682, Latency: 2673ms)

⏳ Waiting 25s for VoyageAI rate limit...

[14/20] How does Jiopay define Director, Nomination and Re...

✅ Grade: A (Quality: 0.801, Latency: 2947ms)

⏳ Waiting 25s for VoyageAI rate limit...

[15/20] What are the Salient Features of the Integrated Om...

✅ Grade: A (Quality: 0.897, Latency: 3833ms)

⏳ Waiting 25s for VoyageAI rate limit...

[16/20] What is the goal of Jiopay ?...

✅ Grade: A (Quality: 0.713, Latency: 2519ms)

⏳ Waiting 25s for VoyageAI rate limit...

[17/20] Where can I download Jiopay?...

✅ Grade: B (Quality: 0.595, Latency: 2709ms)

⏳ Waiting 25s for VoyageAI rate limit...

[18/20] What products are offered by JioPay?...

✅ Grade: B (Quality: 0.685, Latency: 2983ms)

⏳ Waiting 25s for VoyageAI rate limit...

[19/20] I am unable to login to the App/Dashboard. What ca...

✅ Grade: A (Quality: 0.916, Latency: 2726ms)

⏳ Waiting 25s for VoyageAI rate limit...

[20/20] What type of services and merchants are restricted...

✅ Grade: B (Quality: 0.689, Latency: 2730ms)

===========================================================================

🔍 RAG EVALUATION RESULTS - IMPROVED METRICS

===========================================================================

Total Queries: 20 | Successful: 20

📝 ANSWER QUALITY METRICS

Overall Quality: 0.715 (A)

Semantic Similarity: 0.756 (A)

Faithfulness: 0.680

Completeness: 0.717

Relevance: 0.677

Exact Match: 0.469

📊 GRADE DISTRIBUTION

A: ███████████░░░░░░░░░ 55.0%

B: ████████░░░░░░░░░░░░ 40.0%

C: ░░░░░░░░░░░░░░░░░░░░ 0.0%

D: █░░░░░░░░░░░░░░░░░░░ 5.0%

F: ░░░░░░░░░░░░░░░░░░░░ 0.0%

📥 RETRIEVAL METRICS

Similarity Score: 0.756

Threshold Pass Rate: 0.950 (95.0%)

P@1 (improved): 0.743

MRR (improved): 0.756

⚡ PERFORMANCE METRICS

Latency Mean: 2863 ms

Latency P50: 2769 ms

Latency P95: 3480 ms

Memory Usage: 404.7 MB

CPU Usage: 20.3%

💡 INSIGHTS & RECOMMENDATIONS

🟢 Good semantic similarity - answers are well-aligned!

🟡 Moderate latency - acceptable but could be improved

🎯 Strong performance - majority of answers are good quality

⭐ SUMMARY

Overall Grade: A

System Status: 🟢 Excellent

===========================================================================

📁 Results saved to rag\_evaluation\_results.json

🔍 SAMPLE COMPARISONS

--------------------------------------------------

Example 1:

Query: How can I disable SMS notification from dashboard?...

Expected: For disabling SMS Notification you can follow below steps.\nGo to settings tab & select user managem...

Generated: To disable notifications for users, go to the settings tab and select user management [1]. Click on ...

Quality: 0.779 (A)

------------------------------

Example 2:

Query: How long would it require to become P2M merchant after upgradation request?...

Expected: Post Upgradation Request, It would take around 24-48 hours to process the request, subjected to the ...

Generated: After submitting an upgradation request to become a P2M merchant, it will take approximately 24-48 h...

Quality: 0.840 (A)

------------------------------

Example 3:

Query: Why should you consider becoming a part of the JioPay Business Partner program?...

Expected: The JioPay Business Partner Program stands as one of the most rewarding B2B programs in India. As a ...

Generated: You should consider becoming a JioPay Business Partner because it is a rewarding B2B program in Indi...

Quality: 0.892 (A)

------------------------------

(venv) C:\Main Files\Soorya\7thSem\LLM\_Prod\Exp2>