# RAG reliability report: my-rag-v4

Last 7 days, generated 2026-11-19T10:19 UTC

**Healthy.** All SLOs met. Failure rate 0% on 40 evaluated requests.

## Key numbers

| Measure | Value |
|---|---|
| Requests | 40 |
| Evaluated (sampled within budget) | 40 (100.0%) |
| Failure rate | 0.0% |
| Latency p50 / p95 | 2544 / 4276 ms |
| LLM spend / judge spend | $0.00 / $0.03 |
| Abstention quality (mean) | 1.000 |
| Answer correctness (mean) | n/a |
| Answer relevance (mean) | 0.944 |
| Citation correctness (mean) | n/a |
| Context consistency (mean) | 1.000 |
| Context recall (mean) | n/a |
| Context relevance (mean) | 0.787 |
| Faithfulness (mean) | 1.000 |

## Where failures come from

No diagnosed failures in this window.

## Quality changes

- not enough days of data.
- not enough days of data.
- not enough days of data.

## Service-level objectives

| SLO | Target | Actual | Budget left | Status |
|---|---|---|---|---|
| Grounded answers | Faithfulness ≥ 0.7 on 90% | 100.0% | 100.0% | ok |
| Relevant answers | Answer relevance ≥ 0.5 on 85% | 100.0% | 100.0% | ok |

## Content to write

**discount / offer / student** (4 unanswered questions). Write or ingest a document covering: discount, offer, student

**office / tokyo** (3 unanswered questions). Write or ingest a document covering: office, tokyo

Question drift: not enough queries (need 10+ per window)

## Alerts

- [info] [RAGOps] my-rag-v4: SLO 'Grounded answers' recovered (2026-11-19T06:50)
- [warning] [RAGOps] my-rag-v4: SLO 'Grounded answers' at risk (2026-11-19T06:17)