

AI Investment Review Card

Define the outcome. Compare performance. Decide where to invest.

Workflow: _____ Owner: _____ Review period: _____

1

Establish success criteria before measuring

Identify the ideal outcomes of the system.

Customer or business outcome: _____

What counts as success? _____

Quality and risk thresholds: _____

2

Compare like with like

Use comparable workloads and periods; record data sources and sample size.

Measure	Baseline	With AI	Target
Accepted outcomes per period			
End-to-end completion time			
Quality metric or customer impact			
Total cost			
Cost per successful outcome			

Sources / sample / comparison limits: _____

3

Build the cost picture

Include failed attempts and retries. Use the same cost boundary in both periods.

Total cost = AI + tools / infrastructure + human work + rework
Cost per successful outcome = Total cost ÷ Successful outcomes

4

Score the evidence

Assess gains against agreed targets; keep these scores separate.

Outcome gain:	☐ Improved	☐ Unchanged	☐ Worse
Quality and risk:	☐ Pass	☐ Fail	☐ Unknown
Evidence strength:	☐ Strong	☐ Limited	☐ Insufficient

5

Make the next decision

Expand when outcomes justify total cost and quality holds. Test further when evidence is weak.

Current system constraint: _____

Decision: ☐ Expand ☐ Optimize ☐ Test further ☐ Stop

Why this decision: _____

Next intervention + owner: _____

Expected result + review date: _____

AI Investment Review Card

Define the outcome. Compare performance. Decide where to invest.

Workflow	Owner	Review period
Routine customer support resolution	Director of Customer Support	4-week baseline + 4-week AI pilot

1

Establish success criteria before measuring

Identify the ideal outcomes of the system.

Customer or business outcome:
Resolve more routine cases faster and at lower cost while maintaining service quality.

What counts as success?
One case resolved correctly, with no reopening within seven days; escalations excluded.

Quality and risk thresholds:
Accuracy ≥95%; satisfaction ≥4.3/5; no unauthorized data disclosure; human approval for refunds and account changes.

2

Compare like with like

Use comparable workloads and periods; record data sources and sample size.

Measure	Baseline	With AI	Target
Accepted outcomes per period	800	1,000	≥950
End-to-end completion time (median)	12 hours	7 hours	≤8 hours
Quality metric: resolution accuracy	96%	96%	≥95%
Customer impact: satisfaction	4.4 / 5	4.4 / 5	≥4.3 / 5
Total cost	\$24,000	\$23,000	≤\$24,000
Cost per successful outcome	\$30	\$23	≤\$25

Sources / sample / comparison limits: Case records, time logs, invoices; 200 audited resolutions per period. Satisfaction: 160 baseline / 190 pilot responses. Each period: 1,100 incoming cases with similar mix and staffing; seven-day follow-up. Sequential periods may differ.

3

Build the cost picture

Include failed attempts and retries. Use the same cost boundary in both periods.

Total cost = AI + tools / infrastructure + human work + rework
Baseline: \$0 + \$2,000 + \$19,000 + \$3,000 = \$24,000 · With AI: \$1,000 + \$2,500 + \$17,000 + \$2,500 = \$23,000

Cost per successful outcome = Total cost ÷ Successful outcomes → \$30 to \$23 (-23%). Setup: \$12,000, separate.

4

Score the evidence

Assess gains against agreed targets; keep these scores separate.

Outcome gain	● Improved — 25% more accepted outcomes; targets met
Quality and risk	● Pass — thresholds met; no control breaches detected
Evidence strength	● Limited — one team, four weeks; rare risks unproven

5

Make the next decision

Expand when outcomes justify total cost and quality holds. Test further when evidence is weak.

Current system constraint:
Outdated articles drive review effort.

Decision: ● Optimize

Why this decision:
Cost per successful outcome fell 23% with quality held, but outdated articles moved the constraint to human review. Fix it before expanding; evidence is still limited.

Next intervention + owner:
Support Operations Manager improves Context and Memory with current sources and draft citations.

Expected result + review date:
Less review effort and rework; cost per successful outcome below \$23 with accuracy ≥95% and no added retries. Review after 4 weeks + 7-day follow-up.