

WORKING TEMPLATE / NO ASSESSMENT RESULTS

Agree what good looks like.

Use this worksheet with product, engineering, operations and control reviewers before evaluating a financial AI workflow.

1. The workflow and the decision

Workflow: _____

Accountable owner: _____

Decision this evaluation must support: _____

Current system stage and users: _____

2. The baseline

Record the current process and its measurement period. Include task volume, handling time, review effort, errors or exceptions, and the data source behind each measure. Separate observations from estimates.

3. Representative evaluation cases

Define normal tasks, incomplete or conflicting documents, unsupported questions, permission boundaries and expected escalation. Record the dataset owner, permitted use, exclusions and who decides whether a result is correct.

4. People and evidence

Who reviews outputs? What source should accompany an answer? Which actions require human approval? How is a correction recorded? What happens when the system cannot complete the task?

Working note

This is a planning template, not an audit, certification or completed assessment. Thresholds must be set for the actual workflow and its consequences.

WORKING TEMPLATE / DECISION RECORD

Measure, review, then decide.

Complete one record per system version and evaluation run. Keep source evidence beside the conclusion.

5. Acceptance criteria

For each criterion record: measure | test set | acceptance threshold | reviewer.

Answer and citation quality: _____

Permissions and escalation: _____

Latency and cost per task: _____

Operator correction and task completion: _____

6. Failure handling

Record each failure, its severity, affected cases, likely cause and proposed action. Agree what prevents release and what may proceed with a documented constraint. Record the reviewer and decision date.

7. Change and regression checks

Record model, prompt, retrieval, tool and application versions. Identify which tests rerun after a change, how regressions are reviewed and who can approve the resulting release.

8. The decision and the next owner

Decision: proceed / improve and retest / stop.

Evidence supporting the decision: _____

Known limits and actions: _____

Owner and next review date: _____

[Evaluation & Assurance](#)

[Discuss your system](#)