OUTSELL PLAYBOOK

OUTSELL JOURNAL - DECISION WORKSHEET

Enterprise data agents: an evaluation-led change-control model

QUESTION

How should enterprise teams approve changes to a data agent's prompt, model, schema, tools or knowledge sources without relying on production dashboards alone?

GUIDE TAKEAWAY

Treat every agent change as a versioned release. Test a representative task set before exposure, preserve privacy and cost boundaries, watch production drift, and retain a clear rollback decision with named owners.

START HERE

Register the exact agent change and affected boundaries.

YOUR SITUATION

Account or cohort

Decision to make

Starting state and exclusions

Evidence sources, dates, and limitations

Unknowns to validate

WORK THROUGH THE GUIDE

1. Register the exact agent change and affected boundaries

Evidence or input

Output or decision

Owner and review point

2. Freeze a representative evaluation and denial set

Evidence or input

Output or decision

Owner and review point

3. Choose privacy-minimised telemetry and access controls

Evidence or input

Output or decision

Owner and review point

4. Stage rollout with cost, quality and rollback thresholds

Evidence or input

Output or decision

Owner and review point

5. Review drift, gaps and business outcomes by owner

Evidence or input

Output or decision

Owner and review point

COMPARE THE OPTIONS

Aggregate monitoring

Useful when: Volume, latency and spend trends are enough

Limitation: Cannot explain individual failures

Suggested action: Use for low-sensitivity operations

Why this fits our situation

Evidence needed before acting

Governed conversation logging

Useful when: Answer review needs user context

Limitation: Stores potentially sensitive content and has gaps

Suggested action: Limit access, retention and notices

Why this fits our situation

Evidence needed before acting

Full custom traces

Useful when: Tool paths and nested failures need diagnosis

Limitation: Highest collection and governance burden

Suggested action: Instrument only required events

Why this fits our situation

Evidence needed before acting

Offline evaluation

Useful when: A release must be tested before exposure

Limitation: Test sets cannot cover every future input

Suggested action: Combine fixed cases with rotating failures

Why this fits our situation

Evidence needed before acting

COMMIT TO A BOUNDED NEXT STEP

Chosen action

Responsible owner

Required approval

Success measure and definition

Review date

Continue, revise, or stop criteria

Guide: https://outselll.com/blogs/data-agent-change-control-evaluation/

Publisher and evidence policy: https://outselll.com/editorial-policy/