

AI Agent QA Pilot Sample Findings Report

Illustrative template only. This document does not describe a client engagement, contains no customer data, and makes no claimed client results.

PILOT QUESTION	WORKING OUTPUT	DELIVERY POSTURE
Where does a customer-facing AI agent lose reliability?	An evaluation set, error taxonomy, and release recommendation.	Customer-controlled workspace, defined access, documented QA.

1. Review frame

A pilot begins with one bounded workflow. The team agrees the user intents, policy-sensitive moments, escalation paths, and answer-quality standards that matter for that workflow. Reviewers then use a calibrated rubric rather than open-ended opinion.

REVIEW DIMENSION	WHAT IS CHECKED	EVIDENCE RETURNED
Answer quality	Does the response address the user need clearly and completely?	Scored examples with reviewer rationale.
Grounding & retrieval	Is the response supported by the approved knowledge base and the retrieved source?	Citation and retrieval observations.
Handoff & escalation	Does the agent recognize when a human, policy step, or safety check is required?	Escalation gaps and proposed release gates.

2. Illustrative error taxonomy

The labels below are example categories a product team may use. They are not findings from a customer deployment.

CATEGORY	ILLUSTRATIVE SIGNAL	POSSIBLE NEXT ACTION
Unsupported detail	The response adds a fact that is not backed by the approved source.	Tighten grounding rules and add a targeted regression check.
Weak retrieval	The relevant source exists but is not surfaced for the user question.	Review chunking, metadata, ranking, and test queries.
Missed escalation	The agent continues when a handoff or policy review is required.	Clarify threshold language and verify the handoff path.

3. Product-team handoff

WHAT THE TEAM RECEIVES	WHAT IT ENABLES
Calibrated scorecard, labeled evidence, prioritized taxonomy, and an agreed next-release recommendation.	A clear decision on what to fix, what to monitor, and which scenarios belong in the next regression set.