



AI RFP Evaluation Scorecard

Run the same controlled cases through every shortlisted tool. Score observable behavior, not the demo narrative.

PILOT CONTEXT

Tool / edition: _____ Reviewer: _____ Date: _____ Bid case: _____

#	CONTROL CASE	WHAT TO RUN	PASS CONDITION	SCORE 0-3
1	Source authority	Ask where each answer came from and show the exact source.	Source, passage and access boundary are inspectable.	_____ Observed notes on reverse
2	Low confidence	Use an answer that is plausible but not supported by the source set.	The tool flags uncertainty or routes review; it does not guess silently.	_____ Observed notes on reverse
3	Stale evidence	Introduce an approved document that is valid but older than a newer source.	Freshness and applicability remain visible to the reviewer.	_____ Observed notes on reverse
4	Buyer amendment	Change one material requirement after a draft has been reviewed.	Affected answers, evidence and approvals are identified.	_____ Observed notes on reverse
5	Exact candidate	Remove one approved answer from the file selected for release.	The released candidate mismatch is detected before submission.	_____ Observed notes on reverse

SCORING ANCHOR

- 0 Not visible or cannot be demonstrated
- 1 Manual inference or undocumented behavior
- 2 Traceable behavior with explicit human review
- 3 Explicit, controlled and reproducible behavior

DECISION

Total: _____ / 15 Setup time: _____ Correction time: _____

Unresolved gap: _____

Human owner: _____ Next action: _____

COMPARE THE SAME CASES, SOURCES AND RELEASE FILE EVERY TIME.