

Compare evidence, not the sales presentation

A fillable five-page buying aid for teams comparing AI implementation partners, products, or proposals.

Give each provider the same workflow brief, representative permitted examples, expected outputs, system constraints, and acceptance requirements. Record assumptions and exclusions. This is an original Clairvance review method, not an industry benchmark or endorsement.

Rating	Evidence anchor
0 - Missing	No answer, an unsupported claim, or evidence that does not address the requirement.
1 - Described	A general explanation or approach, without a relevant artifact.
2 - Demonstrated	A relevant artifact or demonstration exists, but your agreed requirement is not yet verified.
3 - Verified	Evidence has been checked against your agreed requirement and the result recorded.

Use “open” for missing evidence and “failed” for a confirmed unmet requirement. If an item truly does not apply, agree that before comparing vendors and use the same denominator. No total score can offset a failed critical requirement.

Workflow, systems, evaluation date, and decision owner

Critical requirements that must be satisfied before selection

Save a local copy to enter notes. The file does not transmit answers. Use permitted sample data and role titles; avoid confidential production records in an early demonstration.

Ask for the work behind the demo

Score criteria 1-5 using the anchors on page 1. Record a source for each rating and identify what remains open.

Criterion	Questions and evidence to request
1. Business and workflow fit	Can the provider restate the trigger, outcome, handoffs, and owner? Request a workflow map with assumptions and exclusions.
2. Option judgment	Why this approach rather than a licensed feature, integration, or simpler process? Request an option comparison tied to your actual constraints.
3. System integration	What will connect, through which supported interfaces, and with whose access? Request a data-flow diagram and a dependency list.
4. Representative evaluation	How will ordinary and difficult cases be tested? Request evaluation cases, expected outcomes, review evidence, and retest procedure.
5. Delivery and adoption	What must each team provide? Request milestones, responsibilities, user participation, training, and acceptance decisions.

Ratings 1-5 and evidence references

Missing evidence: request, responsible role, and due date

Use the same examples for every provider. A demonstration tuned to its own convenient examples is not equivalent to an evaluation against your workflow.

Test the difficult operating cases

Score criteria 6-10. Ask the provider to show how the workflow behaves when normal assumptions fail.

Criterion	Questions and evidence to request
6. Data and permissions	Which data can be read, retained, shared, or changed? Request access boundaries, test-data handling, and a stated retention approach.
7. Human review	Who reviews which actions and what context do they see? Request an exception queue example and escalation responsibilities.
8. Failure and recovery	What happens with duplicates, unavailable integrations, and partial completion? Request retry, stop, recovery, and manual fallback demonstrations.
9. Visibility and change	Can the team trace an action and identify a regression? Request logs, monitoring ownership, change records, and retesting rules.
10. Support and handover	Who maintains the workflow after launch? Request documentation, training, support terms, dependency ownership, and exit provisions.

Ratings 6-10 and evidence references

Critical blockers and required evidence to resolve them

NIST AI RMF is voluntary guidance for AI risk management [1]. It does not certify this scorecard or any supplier. Adapt the questions with the responsible technology, privacy, security, and other specialist reviewers.

Know what your team will own

Resolve continuing responsibilities before selecting a provider. Put agreed answers into the relevant project documents.

Area	Buyer questions
Accounts and configuration	Whose accounts run the workflow? Who administers them? Can your team change or revoke access?
Deliverables and ownership	What code, configuration, prompts, rules, documentation, and data are handed over? What remains licensed or restricted?
Recurring dependencies	What subscriptions, model usage, hosting, maintenance, or support remain? How are changes communicated?
Operations and training	Who checks failures and exceptions? What training is included? How is a support request escalated?
Exit and continuity	Can you export information and retain needed records? How is access removed? Can the business continue or use a documented manual route?

Agreed handover deliverables, acceptance evidence, and owners

Contract or operational questions still requiring review

A procurement score is not a substitute for reviewing actual terms. Involve the people who will own the system and any specialists needed for the specific project. Do not accept unsupported claims of certification or guaranteed returns.

Make the decision traceable

Use the ratings to expose differences in evidence. Keep unresolved requirements visible alongside any total.

Comparison	Provider A	Provider B	Provider C
Evidence score	___ / ___	___ / ___	___ / ___
Critical items	Pass / open / fail	Pass / open / fail	Pass / open / fail
Key open evidence	_____	_____	_____
Next step	_____	_____	_____

For 10 applicable criteria, the maximum is 30. Add ratings only after evidence review. If the applicable criteria differ, show the denominator and explanation; do not silently compare raw totals.

Decision: select, request evidence, run a bounded evaluation, or decline

Rationale, unresolved items, and the conditions for proceeding

Decision owner, next action, and review date

Reference and use

[1] National Institute of Standards and Technology. *AI Risk Management Framework*. Overview and resources; accessed September 9, 2026. [Read the official guidance](#).

The questions and scoring anchors are an original Clairvance planning aid, not a validated benchmark. Ratings describe the evidence reviewed, not a promise of future performance.

[Discuss implementation options with Clairvance](#) | Apply the same evidence standard to every provider, including Clairvance.