



CHECKLIST • CHOOSE WHO DOES THE WORK

The AI Partner Scorecard

Twenty questions to ask any consultant, agency, software vendor or partner before you sign, and what a strong answer sounds like.

StrikeArc Labs · October 2026

strikearc.ai/resources/ai-partner-scorecard

The AI Partner Scorecard: twenty questions to ask before you sign.

Every firm that sells AI work sounds confident on a first call. These twenty questions separate the ones who will deliver a working system from the ones who will deliver a slide deck. Ask them of everyone you are considering, us included, and score the answers side by side.

STRIKEARC LABS · OCTOBER 2026 · ABOUT 30 MINUTES PER CALL

The scorecard [PDF](#) The scoring sheet [XLSX](#)

How to use the scorecard.

01 Send the questions ahead.

A firm that knows its answers will put them in writing. Download the PDF and send it before the first meeting.

02 Ask the ones they skipped on the call.

Listen for specifics: a named workflow, a named person, a number with a unit. Vague answers are Partial at best.

03 Mark each answer the same day.

Strong, Partial or Red flag. Mark a question nobody answered as a Red flag; an answer that never arrives is an answer.

04 Compare group by group.

Totals tell you who answered well overall. A red flag on data and access, or on control, outweighs any total, because those are the answers you live with after launch.

Why these twenty.

Most AI projects that disappoint an owner fail in the same places. The first project was picked because it demoed well. Nobody wrote down what the system may touch. Nobody owned it after

launch. The result was counted in activity, and the cost of running it arrived later. Each group of questions covers one of those places.

The five ways to get the work done, from a consultant to an in-house hire, each answer these questions differently. The paper [AI without an AI hire](#) walks through what each one charges for and what you keep. Question 19 has a paper of its own: [measuring AI in dollars](#).

Ask us the same questions.

Open any question in the scorecard to see how we answer it, in the words our site already uses. A few we answer on the call instead, because the answer depends on the play and your systems. Bring the scorecard to a 30-minute intro call and ask all twenty. Compare the answers to those of the other firms you are talking to.

Bring the workflow that costs you the most.

Book a 30-minute intro call at strikearc.ai/book, or write to info@strikearc.ai.

The latest version of this piece lives at strikearc.ai/resources/ai-partner-scorecard.

The AI Partner Scorecard

Mark each answer S (strong), P (partial) or R (red flag). Score S as 2, P as 1 and R as 0, out of 40. A red flag on data and access, or on control, outweighs any total.

QUESTION	CANDIDATE 1	CANDIDATE 2	CANDIDATE 3
The work			
Which of our workflows would you start with, and why that one? Strong: One workflow, named from what you told them, with its volume, what delay or errors cost, and who owns it. Red flag: A platform tour, a list of everything AI could do, or a first project picked before anyone asked about your work.			
What would you tell us not to build? Strong: A written do-not-build list with a reason for each item, and what would change their mind. Red flag: Everything is in scope, and every idea is a good one.			
What do you need from us, and how much of our team's time? Strong: Named roles, the access and examples they need, and the hours, step by step. Red flag: A few workshops to start, then a discovery phase with no end date.			
What will we have at the end that we did not have before? Strong: A working system in the tools your team already uses, running on your data, with a named person responsible for it. Red flag: A report, a roadmap or a prototype that someone still has to build and run.			
Data and access			
What will you read in our systems, and what will you never read? Strong: A written scope, system by system, agreed before anything connects, starting read-only. Red flag: Broad admin access on day one, or no written scope at all.			
Who holds the access you grant, and how do we take it back? Strong: Access you grant per system and can revoke yourself, in one step. Red flag: Shared passwords, or access that only they can switch off.			
What happens to our data when we stop working together? Strong: A written retention policy with numbers: what is deleted, when, and how you ask for it. Red flag: No policy, or data kept indefinitely to improve their product.			
Will what you learn in our business be used for other companies? Strong: A plain answer in writing, before any data moves. Red flag: A hedge, or an answer that only exists in the terms of service.			
Control			

QUESTION	CANDIDATE 1	CANDIDATE 2	CANDIDATE 3
Can anything send, pay, post or change a record without a named person approving it? Strong: A list of what the system may do on its own under your rules, and what waits for a named approver. Red flag: Full autonomy offered as the selling point.			
Who approves a change to how the system behaves? Strong: A change you can read before it goes live, with a named approver on your side. Red flag: Updates you find out about afterwards.			
What does the system record, and can we read it? Strong: Every action on the record, readable by your team. Red flag: A system nobody on your side can inspect.			
What happens when the system is unsure? Strong: It hands the item to a person, with the evidence attached. Red flag: It makes its best guess and moves on.			
After launch			
Who runs the system after launch, and what does that cost? Strong: A named person or team, and a stated price that includes running it. Red flag: Your team takes it from here, with no handover plan, or a support contract that appears later.			
How will you know it has broken or drifted before we do? Strong: It reports on a schedule, a silent system raises an alert, and quality is checked on your own examples. Red flag: Call us if something looks wrong.			
What happens when the AI models change or get cheaper? Strong: They re-test on your examples before switching, tell you, and pass the saving through. Red flag: No plan, or you find out from the bill.			
What do we keep if we stop? Strong: Clear terms on your data, your configuration and your documents, and how the handover works. Red flag: Everything stays with them, or nobody can say.			
Money and proof			
How do you charge, and what is included? Strong: A stated price, what it covers, and when billing starts. Red flag: Open-ended hours, or no price even for work they have done many times before.			
How long do we commit for, and how do we leave? Strong: A stated term and a stated way out, in the proposal. Red flag: A long lock-in, or a renewal buried in the contract.			

QUESTION	CANDIDATE 1	CANDIDATE 2	CANDIDATE 3
How will we measure what it returned, and who signs the number? Strong: A baseline taken before the work starts, results in dollars or hours, and a named person who signs each figure. Red flag: Activity counted as results: messages sent, tasks automated, hours of AI usage.			
What results can you show us, and where do the numbers come from? Strong: An honest answer either way: results they can name, or worked examples clearly labelled as illustrative, and an offer to measure yours. Red flag: Return figures with no company, no method and no label.			
Total, out of 40			