The AI vendor claim checklist

26 questions to ask before you believe what an AI vendor says about accuracy, automation, your data, cost and lock-in. Take it into the demo call.

> **How to use it.** Send the questions to the vendor before the demo. Tick each one only when you have a specific, written answer; a vague answer doesn't count. Anything unticked at the end is a risk you are agreeing to carry.

1. The claim itself

Before judging whether a claim is true, pin down what is actually being claimed.

- [ ] **Can the claim be written as one specific sentence with a number, a unit and a timeframe?**
  "Cuts support tickets by 40% within 90 days" can be checked. "Transforms customer support" can't.
- [ ] **Is it measured, reported, a forecast, or the vendor's opinion?**
  A benchmark result, a customer's account, a prediction and a slogan deserve very different levels of trust.
- [ ] **Compared with what?**
  Faster than a human, than the old version, than doing nothing, or than a competitor? A number without a baseline means nothing.
- [ ] **Who measured it, and are they independent of the vendor?**
  Vendor-run tests, paid analyst reports and customer case studies written by the vendor's marketing team are not independent.

2. Accuracy and benchmarks

Accuracy figures are only meaningful with their conditions attached.

- [ ] **Which benchmark or dataset produced the number, and does it look like your data?**
  A model that scores well on a public benchmark can still fail on your documents, language or edge cases.
- [ ] **What was the sample size, the date of the test, and the exact model version?**
  Results from 50 examples, or from a model version that has since changed, don't describe what you'd be buying today.
- [ ] **How often is it wrong, what do the errors look like, and how are they shown to the user?**
  A 95% success rate matters less than whether the other 5% fail loudly or slip through as confident, plausible answers.
- [ ] **Can you test it on your own data, against success criteria you set, before you sign?**
  A vendor that won't allow a trial on real data is asking you to trust the demo.

3. “Agent” and automation claims

When a product acts on its own, the questions shift from accuracy to control.

- [ ] **Which steps run with no human in the loop?**
  “Autonomous” can mean anything from drafting a reply for approval to sending it, refunding money or changing records.
- [ ] **What access does it need, and can it be limited to read-only or to specific records?**
  Broad permissions turn one bad output, or one malicious instruction hidden in a document, into real damage.
- [ ] **Is every action logged so you can see what it did, when, and why?**
  Without an audit trail you can't investigate a mistake, prove compliance, or improve the setup.
- [ ] **How do you pause it, and can its actions be undone?**
  Ask to see the off switch and the rollback in the demo, not just the happy path.

4. Data, privacy and security

AI features often send your data somewhere new. Find out where before it goes.

- [ ] **Is your data used to train the vendor's models, or anyone else's? Is that off by default?**
  Opt-out settings buried in an admin page are easy to miss; the contract should say it plainly.
- [ ] **Where is your data processed and stored, for how long, and how is it deleted?**
  Regions, retention periods and deletion on request matter for GDPR and most customer contracts.
- [ ] **Which outside model providers and subprocessors handle your data?**
  Many AI features call a third-party model API. That provider's terms apply to your data too.
- [ ] **Do their security reports (SOC 2 Type II, ISO 27001) actually cover the AI feature?**
  A certificate for the core product doesn't automatically cover a newly added AI service.
- [ ] **How have they tested against prompt injection and data leaks between customers?**
  Instructions hidden in an email, web page or file can make an AI tool reveal or do things it shouldn't.

5. Cost

AI pricing is often usage-based, which makes the demo price a poor guide to the real bill.

- [ ] **What is the pricing unit (seats, usage, tokens, “credits”), and what happens at ten times your expected volume?**
  Credits and tokens are hard to forecast. Ask for a worked example at your real volume.
- [ ] **Is every feature you saw in the demo included in the plan you're being quoted?**
  The most impressive features are often in a higher tier or a paid add-on.
- [ ] **Can you set a hard spending limit, and are overages capped or billed automatically?**
  A runaway workflow should hit a ceiling, not your card.
- [ ] **What will setup, integration and human review of outputs cost in staff time?**
  Time spent checking and fixing outputs is part of the real cost, and it often decides whether the tool pays off.

6. Lock-in and change

Models change under the hood. Make sure you can live with that, or leave.

- [ ] **Can you export your data, settings, prompts and outputs in a standard format?**
  If leaving means rebuilding everything by hand, you have less leverage at renewal.
- [ ] **Will you be told when the underlying model changes, and can you stay on a tested version?**
  A silent model update can change results you've already validated.
- [ ] **What notice do you get for price or terms changes, and what is the exit clause?**
  Fast-moving AI products reprice often. Long contracts without protection can lock in a bad deal.

7. Proof from customers

Case studies are marketing. References you choose and call yourself are evidence.

- [ ] **Can you speak to two customers of your size and industry, picked from a list rather than handed to you?**
  Ask them what didn't work, what it really costs, and whether they'd buy it again.
- [ ] **For each case-study number: over what period, under what conditions, and who wrote it?**
  “Up to 10x faster” over two weeks in one team at a launch customer is not a result you should plan around.

Red flags

- “Up to” numbers with no average, baseline or sample size.
- Accuracy quoted without saying what was tested or when.
- “AI-powered” with no explanation of what the AI actually does.
- No trial on your own data, or a trial only on the vendor's sample data.
- Only testimonials and logos, no references you can call.
- Vague answers about where your data goes or whether it trains models.
- Pressure to sign before a model or pricing change “ends the offer”.

Free to use and share. If you quote it, please link to getcoretech.net/resources/ai-vendor-claim-checklist · © GetCoreTech