AI solution due diligence guide for accounting firms

A practical evaluation framework for AI-enabled technology solutions.

Developed by the CPA.com AI Working Group

In collaboration with RADICAL+

Executive summary

As AI becomes increasingly integrated into accounting and professional services, selecting the right solution partner is critical, not just for efficiency, but for maintaining the trust and security that defines the profession. This guide is designed to help you have informed, productive conversations with AI vendors and solution providers.

These aren't "gotcha" questions. They're meant to help you understand how a solution works, where your data goes, and whether a provider has thought through the complexities that matter in regulated environments. Strong vendors will welcome these conversations as they're an opportunity to demonstrate their commitment to providing AI responsibly.

Use this as a framework for your evaluation process. Not every question will apply to every solution, and you may have additional concerns specific to your practice. The goal is to ensure you're making decisions with clarity and confidence.

The document is broken into 3 sections:

1. Quick start: Top 5 things to verify first
2. Expansive questions to ask vendors
3. Due diligence evidence checklist

Quick start:
Top 5 things to understand

1

Where does your data go?

Request a simple data-flow explanation showing:

- What data is accessed
- Where data is processed and stored
- Whether data is sent to an external LLM (e.g., OpenAI, Anthropic, etc.)
- What is redacted or anonymized before leaving the system

Vendors should be able to provide this quickly and clearly. Note, it is important to understand the agreement between the vendor and the LLM operator (which may or may not be the LLM developer). This agreement will cover elements of data privacy and, importantly, whether the operator uses customer data to train or otherwise improve the model.

2

Where and why is AI used inside the product?

Ask: **"Where do you use generative AI — and why in those areas?"**

This distinguishes between:

- Deterministic logic (used where accuracy must be exact)
- Probabilistic AI (predictive)
- Generative AI
- Agents or autonomous workflows
- Other types of AI

3

SOC 2

Not all SOC 2 reports are equal. Ask if the vendor has a SOC 2 report available and request a copy.

- Type 2 (preferred as it covers a period of time vs. a point in time)
- Scope covers the product you're buying
- Overall opinion and exceptions and how they were resolved
- What subservice organizations are utilized and if they are excluded ("carved out") of the scope of the report
- Auditor reputation

CPA.com

4 Trial period & validation testing with your own data

Ask for a trial of any solution you are vetting. Every firm should test accuracy, behavior and hallucinations using firm-specific or anonymized data. Be sure to understand what happens with your data after the trial period before uploading anything with client or sensitive information.

5 Vendor transparency

Look for:

- Clear, complete answers
- Willingness to provide artifacts (SOC report, sub-processor lists, etc.)
- Clarity about types of AI usage (i.e., models and architecture)
- No vague or evasive answers (i.e., shipping full data to an external LLM is a red flag)

Expansive questions to ask vendors

If you want to dig deeper into a specific area, here are some suggested, more advanced questions that you can ask. Remember, depending on the type of solution and the stage of the company, the vendor may not have all of the requested information from this section.

Data privacy, ownership & use

This section helps you understand what happens to your firm's and clients' data: where it lives, how it's used and what control you retain over it.

What data will your system access and store?

Where is my client's data stored and for how long? What is your data retention and deletion policy?
Context: Look for vendors who can clearly describe their deletion processes, export options, and how they handle data from test or pilot phases (not just a general "we delete it when you ask" response).

What happens to our data if we discontinue service?

Will my data be used to train your models? Do we have the ability to opt out of data being used for training or model improvement?
Context: A credible vendor will explain training boundaries, opt out mechanisms and how prompts are handled. Be cautious of vague assurances like "we don't train on your data" without supporting detail about how that's enforced.

Will my data be sent to any third-party models or providers? If so, what data and how is it redacted?
Context: The maturity of a vendor's redaction process, how they mask names, PII and sensitive fields before making model calls is often a good indicator of their engineering sophistication.

Who owns intellectual property created from my data?

Can we export our data and logs if we move platforms?

Where do you use generative AI — and why?
Context: This helps firms distinguish between deterministic logic (preferred where accuracy is required) and generative tasks.

Do you automatically enable AI features, or are they opt-in? If automatically enabled, how do you notify customers?

Auditability, explainability & controls

This section focuses on transparency and the ability to understand, explain and audit AI behavior.

Can we review an audit trail of AI actions?

Are AI-assisted decisions explainable and traceable for compliance and regulatory requirements?
Context: "Black-box AI" may not be acceptable in regulated environments.

Can we export audit logs if needed?

How do you handle audit requirements for overrides or adjustments?

Security & compliance

This section addresses how the vendor protects your data and meets regulatory requirements — essential for maintaining client trust and professional obligations.

Do you have a SOC 2 report, and can we have a copy?

Do you comply with GDPR and other applicable privacy regulations?

Who within your organization can access client data, and how is access controlled?

How do you enforce least-privilege access to client data?
Context: Least-privilege means users and AI systems can only access information required for their specific role. The vendor should explain how they restrict access to sensitive data, ensure AI cannot "see" more than a human user is authorized to see, and prevent engineering or support staff from having broad access by default. Answers like "only trusted people can access it" or "we manually control access" may indicate weak controls.

Do you inherit and enforce our firm's access permissions and role-based controls for AI agents?
Context: The AI should not see more than the user is permitted to see.

How do you ensure AI cannot bypass user permissions or access data outside a user's scope?

Do you provide detailed audit trails for all system and AI-generated actions?
Context: If audit logs are "coming soon," that signals the system may not be ready for production use.

Can you provide a list of sub-processors and what data they receive?

Can you share any historical incident reports or downtime summaries for transparency?

Architecture, models & technical approach

This section helps you understand which AI models are being used, how they're deployed, and whether the approach aligns to the use case.

What AI models are you using, and why did you select them for the specific use cases?
Context: Vendors should distinguish between generative models, predictive algorithms, and deterministic logic.

How are those models hosted (your infrastructure, private cloud or external API)?

How do you handle new model releases?
Context: Look for evidence of versioning, testing and controlled rollout processes, not "we swap it in when it's available."

How do you isolate and secure agent access within your system?

What safeguards exist to prevent unintended access or data leakage?

Is any part of your product primarily a wrapper on top of GPT or another LLM?
i.e., Does the vendor add significant licensed proprietary context (RAG) to the LLM, or is the prompt sent largely "as is"?

Can you provide a data-flow diagram showing how data moves through your system?

© 2025 CPA.com. All rights reserved

Accuracy, reliability & quality controls

This section examines how the vendor ensures their AI produces reliable, trustworthy results — and what happens when it doesn't.

How do you measure model accuracy, and what benchmark or dataset is used?
Context: Be wary of anecdotal claims like "98% accuracy" without understanding the underlying methodology.

How do you detect low-confidence results, and how are they surfaced to users?

What happens when the AI is unsure or unable to confidently complete a task?
Context: There should be escalation paths or human-review triggers.

Do you provide confidence scoring and transparency into when human review is required?

How do user corrections or feedback improve the system?
Context: Look for structured feedback mechanisms, not vague claims that "the AI learns on its own."

How do you ensure consistency and avoid hallucinations?

Do you use deterministic logic where appropriate?
Context: Not everything should be generative AI-driven.

Can you provide anonymized accuracy benchmarking across 20+ customers?

Will you support validation testing during our trial period?

Risk, governance & vendor maturity

This section assesses the vendor's operational maturity and their ability to support long-term partnership.

What human-in-the-loop review processes exist?

How do you monitor for bias, drift and unexpected outputs?

What is your process for model governance and review?

How quickly can the system be rolled back if issues arise?

What are your cyber-incident and AI-related breach protocols?

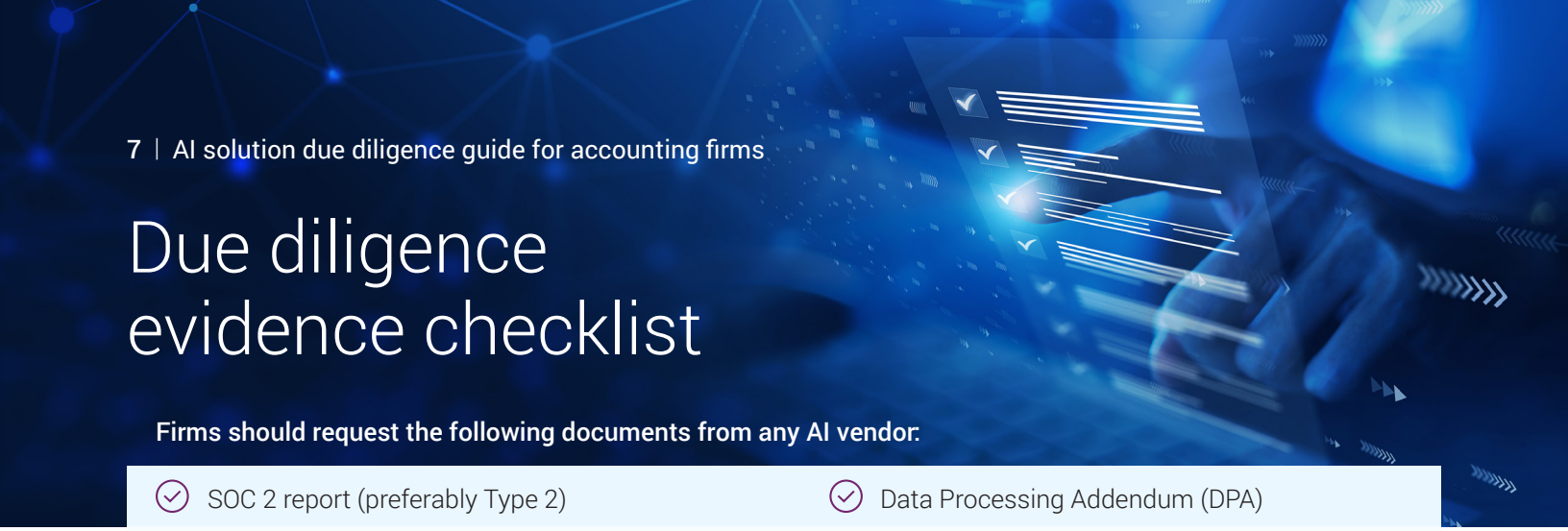
What deployment support, training and change management assistance do you provide?

How do you evaluate and prioritize new use cases and enhancements?

How long has this AI system been deployed and in production?

Can you demonstrate customer references for similar firm sizes and use cases?

How do you ensure continuity if the underlying model provider changes or sunsets a capability?



Due diligence evidence checklist

Firms should request the following documents from any AI vendor:

- ☑ SOC 2 report (preferably Type 2)
- ☑ Service Level Agreement (SLA)
- ☑ User agreement
- ☑ Sub-processor list
- ☑ GDPR/privacy policy
- ☑ Data Processing Addendum (DPA)
- ☑ Data-flow architecture diagram
- ☑ Incident response plan
- ☑ Historical incident reports (if available)
- ☑ Accuracy benchmark summary

How to review a SOC 2 report

Check the scope
Does it cover the product *you* are buying?

Check the type
Type 2 shows controls were tested over time.

Check the audit period
Ensure recency and continuity (i.e. there are not gaps in the period of time between reports).

Check exceptions
Vendor should explain all exceptions and how they are being addressed.

Check the auditor
Evaluate the reputation of the auditor and rigor of the report itself.

The AICPA has published helpful resources that dive into this topic in more depth.

Evaluating AI add-ons from existing vendors

Remember that vendors you have been working with for years are adding new AI features regularly. It is important that firms treat new AI features as separate products and perform due diligence on existing vendor relationships regularly, including but not limited to:

- Request updated data-flow documentation
- Validate security and redaction changes
- Confirm opt-in vs. automatic enablement
- Confirm no new sub-processors were added
- Request updated Terms of Service
- Conduct a trial period (not just trust marketing)

CPA.com

Using this guide

Think of these questions as conversation starters, not a checklist to race through. The quality of the answers, and the vendor's willingness to engage transparently, will tell you as much as the answers themselves. Strong partnerships are built on shared understanding and trust. The right vendor will appreciate your diligence and see it as a sign that you take your professional responsibilities seriously. This resource is not an exhaustive list, and depending on your organization's policies and/or client regulatory requirements, additional considerations may be required.

The CPA.com AI Working Group created this resource to help firms make confident, informed and forward-looking decisions as they integrate AI into their practices. Thank you to the working group for their continued contributions and insights

CPA.com