taxgpt

The Agentic Tax Firm

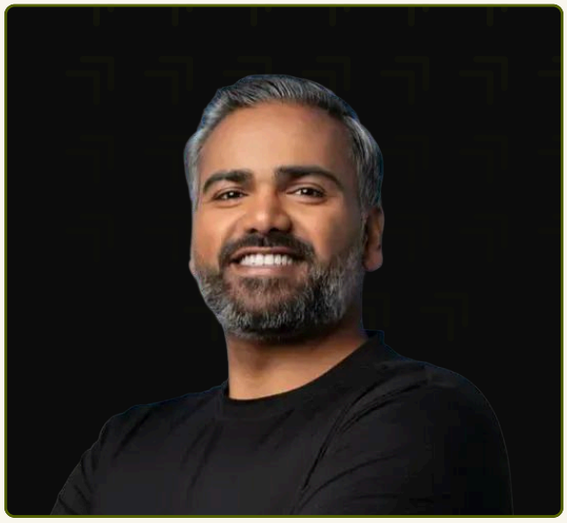
How Forward-Looking Practices Are Restructuring Work Around AI

Table Of Contents

Author's Note

I am the founder of TaxGPT, an AI-powered tax research and communication platform. While I reference TaxGPT throughout this handbook as an example I know well, the principles and strategies discussed apply across many AI solutions. I encourage you to evaluate multiple tools and choose those that best fit your firm's specific needs.

Kashif Ali
CEO & Founder



The tax profession is not waiting for AI to arrive. **It is already sorting into two groups:** firms that are restructuring around what agentic AI can do, and firms that are still treating it as a research accelerator while the distance between those two positions grows. This document is for principals who want to understand that gap clearly, assess honestly where their firm sits, and know what the first steps toward closing it actually look like.

What "Agentic AI" Actually Means (and Why It's Different This Time)

Most tax professionals have been using AI for a couple of years now. The way it has been used is you ask it a research question and it gives you an answer. Sometimes you might paste a statute and it would summarize the relevant provisions. You describe a client situation and it will give you a rough draft for a memo response. Although this model of AI is useful, **there is a ceiling where it can only respond and not act.**

Agentic AI refers to systems that can execute a sequence of tasks toward a defined goal without needing a prompt at every step. Rather than waiting for input, an agent receives an objective, determines what steps are required to complete it, uses the tools available to it, evaluates what comes back, and continues until the work is done. The practical significance of that for tax work is not subtle. A research tool tells you what the IRC says about a specific provision. An agent pulls the client's documents, identifies which provisions apply to their situation, runs the relevant research, drafts the analysis, and flags the judgment calls that need a partner's attention. The first tool speeds up one step in a workflow. The second completes the workflow.

Understanding where human oversight fits into that picture matters because firms are implementing agentic AI across a wide range of configurations. In some workflows, a professional remains closely involved throughout, reviewing outputs at each stage before the system proceeds. In others, the agent handles the full sequence end to end and surfaces only the decisions that genuinely require professional judgment. Neither model is inherently right. The right configuration depends on the workflow, the risk level, and how much confidence a firm has built in the system's outputs for that specific task. What matters is that firms are making that choice deliberately rather than defaulting to the lowest level of automation because it feels safer.

The distinction between AI as a tool you use and AI as a participant in your workflows is where the operating leverage actually lives. Firms that have crossed that line are not just completing the same work faster. They are completing more of it, with fewer people involved in the parts that do not require human expertise, and redeploying that capacity toward the advisory work that clients pay for. So now let's break down how that operation model looks like.

The Firm Operating Model That’s Breaking Down

The traditional tax firm is built as a pyramid, and for most of the profession's history that shape made sense. Work enters at the bottom, where staff and seniors handle the volume tasks: gathering documents, entering data, running initial calculations, assembling returns. It moves up through managers who review for accuracy and flag issues, and arrives at the partner level for final sign-off and client communication. Each layer adds a judgment call. It also adds time, because no step can begin until the one below it is finished.

That model was designed around the constraints of manual work, and it optimized well for those constraints for decades. The problem is that the constraints it was built around are no longer fixed. The volume tasks at the base of the pyramid, the ones that justify the staffing structure above them, are exactly the tasks that agentic AI handles most effectively. When document intake, data entry, initial research, and return assembly can be completed by an AI system in a fraction of the time it takes a staff accountant, the economic logic of the pyramid starts to come apart. You do not need the same base to support the same top.

The review bottleneck at the partner level is where this breaks down most visibly. In a typical return cycle, a significant share of total time is spent not on the work itself but on the handoffs between levels — the wait for a senior to review what staff prepared, the wait for a manager to clear what the senior flagged, the wait for a partner to sign off on what the manager escalated. Each of those transitions introduces delay, and during peak season those delays compound. A firm running at capacity is often not limited by how fast its partners can work but by how fast work can get to them in a reviewable state. Agentic AI does not just accelerate the steps in that chain. It compresses the chain itself by handling the assembly work that creates the queue.

The billing structure that evolved alongside the pyramid model adds a second layer of pressure. **Hourly billing made sense when hours were the primary input and expertise was distributed unevenly across the team.** Clients are increasingly resistant to that model, partly because they can see that AI is changing how long things take, and partly because the most valuable thing a tax firm delivers has never actually been hours. It has been judgment, risk management, and advice that prevents problems rather than just reporting them. Firms that restructure around agentic AI find themselves able to move toward value-based pricing not because they decided to, but because their cost structure no longer requires them to bill for the time AI now handles.

The gap between what clients expect and what traditionally-structured firms can deliver is widening on a predictable trajectory. Clients who have seen AI compress timelines in other professional services are bringing those expectations into their tax relationships. They want faster turnaround, more proactive communication, and advisors who have time to think rather than process. A firm still running a full manual workflow at the base of its pyramid will find that gap increasingly difficult to explain. The firms that are closing it are not doing so by working harder within the existing structure; they are figuring out which workflows can be optimized with agentic AI.

The Three Workflows Agentic AI is Restructuring Right Now

The clearest way to understand what agentic AI changes in a tax firm is not to look at the technology but at the **workflows** it touches. Three workflows are being restructured right now across practices of every size, and each one follows the same pattern: a process that required human involvement at every step is becoming a process where human involvement is concentrated at the steps that actually require judgment.

Tax Research

Research in a traditional firm workflow is iterative by necessity. A practitioner identifies the question, searches the relevant primary sources, evaluates what they find, determines whether it answers the question or surfaces a related issue that needs to be resolved first, and continues until they have a defensible position. That process is not inefficient because practitioners are slow. It is time consuming because it requires sequencing, and each step depends on evaluating the output of the last one.

Agentic AI handles that sequencing autonomously. A research agent receives the question, identifies the relevant statutory provisions, retrieves authoritative sources across the applicable jurisdictions, evaluates whether the authorities are consistent or in conflict, and returns a structured analysis with citations the practitioner can verify and use. What previously took two to four hours for a complex multi-issue question now takes minutes, with the practitioner's time concentrated on evaluating the conclusion rather than building toward it.

Return Prepared

The preparation stage of a return cycle is where volume labor is most concentrated and where the connection between time spent and value delivered is weakest. Document collection, data extraction, cross-referencing prior-year returns, populating forms, and ensuring consistency across schedules are tasks that require accuracy and attention.

Agentic return preparation systems are now starting to handle this workflow end to end. They pull source documents from wherever they live, extract the relevant data, map it to the correct form fields, and integrate directly with the tax software the firm already uses, without manual data entry and without requiring the firm to change its existing workflow infrastructure. The preparer who previously spent hours on a 1040 with a Schedule C and a rental property is now reviewing a completed draft in a fraction of that time, with attention focused on the judgment calls rather than the assembly work that preceded them.

Return Review

Reviews have always been the bottleneck which in part is because the review process itself is unstructured. A reviewer scanning a completed return is simultaneously checking for transcription errors, evaluating classification decisions, assessing audit exposure, and identifying planning opportunities the preparer may have missed. They are doing all of this in a single pass through a document that contains dozens of potential issues. The cognitive load is high and the process does not scale.

Agentic review systems work through that same document systematically, flagging issues by category before a human reviewer ever opens the file. Transcription errors and omissions surface as one category of finding. Potential audit triggers surface as another. Missed savings — depreciation elections not taken, retirement contributions not maximized — surface as a third. By the time a partner opens the file, the mechanical work is done. What is left is the judgment: whether the positions are supportable, whether the client's situation warrants a different approach, whether anything needs a conversation. That is what partner time is actually worth, and it is rarely what partner time is spent on during a busy season.

Where This Is Heading

Research, preparation, and review represent the core of the return cycle, and restructuring those three workflows alone changes the economics of a tax practice substantially. They are not, however, the boundaries of what agentic AI will touch. The workflows adjacent to the return cycle — client communication, bookkeeping, audit support, engagement management — are on the same trajectory. The firms building operational fluency with agentic AI now, will be positioned to absorb those expansions as they arrive. There are some considerations to take into account though before deep diving into agentic AI and we cover that in our next section.

What Separates Purpose-Built Tax AI From General AI

Every CPA who has experimented with general-purpose AI for tax work has eventually encountered the same problem: a confident, well-structured answer that is wrong. **Not vague, not incomplete — wrong, with citations that either do not exist or do not say what the AI claims they say.** For a practitioner whose professional liability attaches to the advice they deliver, a wrong answer stated confidently is not a minor inconvenience. It is a risk that the practitioner has to catch before it reaches a client, which means the time savings the tool appeared to offer were never real.

The reason general-purpose AI fails in this specific way comes down to how it was built and what it was built to do. A large language model trained on broad internet data is optimized to produce text that is coherent, fluent, and contextually appropriate. It is not optimized to distinguish between a Revenue Ruling and a blog post summarizing a Revenue Ruling, and it does not understand that one constitutes primary authority and the other does not. It has no concept of jurisdiction hierarchy, no awareness that a Tax Court memorandum decision carries different weight than a published opinion, and no mechanism for recognizing when a regulatory position it learned during training has since been updated or reversed. **When it cites an IRC section or a Treasury Regulation, it is pattern matching on what a citation looks like, not retrieving a verified source.**

Purpose-built tax AI starts from a different foundation. The training corpus is not the internet broadly but a curated set of authoritative sources: the IRC, Treasury Regulations, IRS guidance, court decisions, revenue rulings, and provincial and federal tax law across applicable jurisdictions. Every source is vetted against its primary authority, and the system is built to distinguish between authoritative guidance and secondary commentary. When a purpose-built system cites a source, that source exists, says what the citation claims it says, and has been verified against the underlying primary authority. That distinction is not a product differentiator in the marketing sense. It is the difference between a work product a practitioner can defend and a work product that creates liability.

The practical question for any firm evaluating AI tools is not whether AI can help with tax work, but whether the AI being used understands tax well enough that its outputs can be trusted as a starting point rather than treated as a first draft that requires independent verification of every claim. A tool that requires a practitioner to re-research everything it produces is not saving time, it adds a secondary verification step. Purpose-built tax AI earns a different level of trust because it was trained to meet the standard the profession requires, not adapted to approximate it.

Next we will explore this Agentic Firm Maturity model.

The Competitive Pressure That Isn't Waiting

The tax profession has absorbed technology changes before, and the historical pattern has generally been that adoption spreads gradually enough that no single firm gains a durable advantage from moving early. That pattern is not holding this time, and the reason is the magnitude of the capacity change involved. Previous technology shifts, practice management software, digital document handling, cloud-based filing tools, made existing workflows faster. Agentic AI changes which workflows require human time at all. That is a different order of change, and it is **compressing the window between early adoption and table stakes faster than most firm principals have accounted for.**

The capacity gap is the most immediate competitive pressure. A firm that has restructured its return preparation and review workflows around agentic AI can handle materially more returns with the same headcount, or the same volume with a leaner team. That additional capacity means taking on clients that a traditionally structured competitor simply cannot serve without first hiring more staff. It also means pricing those engagements more competitively, because the cost structure underneath them has fundamentally changed. This advantage is not marginal. Firms that have made the transition are not 10 or 15 percent more efficient on the workflows AI touches. The efficiency differential on specific tasks is an order of magnitude larger, and **it flows directly into client capacity and margin.**

The talent dynamic is shifting alongside the capacity dynamic, and the two are connected. When volume work is handled by AI systems, the practitioners a firm needs are the ones who are strong at judgment, client relationships, and complex advisory work, not the ones who are fast at data entry and return assembly. That changes what a firm looks for when hiring and what it can offer in terms of the work people actually do day to day. Firms that have made the transition are finding it easier to attract experienced practitioners who want to spend their time on advisory work rather than production, **and harder to justify roles structured primarily around tasks that AI now handles more reliably.** The staffing model and the talent market are both adjusting to the same underlying shift in what the work actually requires.

The billing model pressure is the third force, and it is being driven partly by clients rather than entirely by internal firm economics. Clients who have watched AI compress timelines in legal services, in financial analysis, in software development, are increasingly aware that the hourly billing model in tax reflects legacy cost structures more than it reflects the value being delivered. The firms that get ahead of that conversation are the ones restructuring their pricing around outcomes and advisory value rather than waiting until clients raise the question directly. The move from hourly to value-based billing is not straightforward, and it requires the kind of confidence in delivery quality that only comes from having reliable systems underneath the work. But the direction of travel is clear, and the firms that have already restructured their workflows are better positioned to make that pricing transition than those still building the case internally for why it is necessary.

What To Do With This

Understanding where the profession is heading and understanding what your firm needs to do next are two different things, and most principals who have read this far are already thinking about the second question. The transition to an agentic operating model does not require a firm-wide transformation initiative. It requires starting in the right place with enough structure to learn from the first 30 days before expanding further.

Start with an honest audit of where hours currently go — not at the firm level, but at the workflow level by task type. Most firms that run this exercise find the distribution more skewed toward volume work than expected. That skew identifies where the agentic opportunity is largest and which workflow to prioritize first.

Once that workflow is identified, run a parallel test. Put one return type or one category of research question through an agentic AI system alongside your existing process for 30 days without changing anything else. Let the results tell you what level of oversight is appropriate before you build the new process around it. By 90 days in, a firm that has run a disciplined test on one workflow typically has enough confidence to expand to a second.

TaxGPT gives tax firms the tools to run that test today.

Tax research, return preparation through Cowork, and return review through Agent Andrew are purpose-built for tax, citation-backed, and integrated with the software your firm already uses. If your firm is ready to identify the first workflow and start, a free trial gives you full access to all three.

Here Is Our Personal Commitment To Security

Ready to Transform Your Tax Practice?

Join the thousands of CPAs and tax professionals who are working smarter with TaxGPT

Why Leading Tax Professionals Choose TaxGPT:

Save Billable Hours 01

Reduce research time from hours to minutes while maintaining the technical rigor your client expects.

Enhance Client Service 02

Deliver faster, more comprehensive responses backed by authoritative tax guidance.

Stay Current Effortlessly 03

Access up-to-date tax law changes, rulings, and technical guidance without manual research.

Protect Your Practice 04

Work within a secure platform designed specifically for sensitive tax and financial information.

Get Started Today

Schedule A Demo

taxgpt