"AI software development is completely different" - how Sage plans to deliver enterprise AI with thousands of customer-specific finance models

By Jon Reed March 29, 2024

I expected to hear a well-thought AI strategy at Sage Transform. What I didn't expect were the surprises, including thousands of customer-specific finance models - and a mysterious non-product called the Sage AI Factory. Here's what I learned on the ground, including a sit down with Sage CTO Aaron Harris.

(Aaron Harris, Sage CTO at Sage Transform 2024)

My biggest frustration of the spring event season? The dreaded AI news chase. One great thing about the SaaS vendor emergence: most of the new functionality from the keynote stage was already shipping. "*You can go back to your project, and turn this functionality on now*" - what a winning keynote line!

Alas, that's all but forgotten now, with vendors racing to stake their gen AI claim, and pushing out AI previews. Meanwhile, customers have immediate questions about release timeframes, pricing, and data movement. Thus the on-the-ground AI news chase, as your trusty diginomica team hustles for roadmap specifics.

But there was a welcome highlight of the spring events so far. Just about every enterprise software vendor has a different approach to so-called "responsible AI," with AI architectures designed to reduce hallucinations and increase output relevance. Sage was no exception - and Sage Transform brought AI surprises I wasn't expecting.

Sage's AI plans bring surprises - thousands of finance models?

It all started a week before the event, during an online analyst session. Sage CTO Aaron Harris basically said Sage *could provide custom models to any customers on any release.* Customers did not have to be running on Sage Intacct's multi-tenant apps.

The Sage Network is designed to enable all Sage customers to consume cloud services from any release, without any forced march ERP upgrades. The same is true for AI services. But hold up - from a technical standpoint, vendors have told me deploying AI is much easier for their SaaS/cloud customers, compared to on-prem, where customized landscapes complicate the issue. *So what has Sage figured out?*

There's a fair amount to break out here, but here's the short version:

1. Established components of Sage AI, such as Accounts Payable Automation and GL Outlier Detection, are already live and in play.
2. A Sage Copilot digital assistant for finance, demoed at Sage Transform, is still under development.
3. A partnership with AWS for what Sage describes as the first domain-specific accounting Large Language Model (LLM). Harris says Sage's plans also include "*tens of thousands of models that we've deployed to support our customers, that are retrained on a daily basis.*"
4. Delivery of AI services happen through the Sage Network, via a framework Harris informally refers to as the "AI Factory." The AI Factory is not a Sage product offering per se. It's the engine that enables AI via customized models to any Sage customer, utilizing their own data, with all the security precautions such a framework requires.

For point number one - and more context on Copilot - check my prior piece, Sage Transform 2024 - "end the monthly close" raises the stakes - but does AI bring real-time finance within reach? I also explored customer reaction to the Copilot demo in Sage Transform 2024 - ACT Construction shows why industries are the future for cloud ERP.

So what about those finance-specific language models? Enterprise software vendors face a defining moment: *how much AI do you build, and what do you take off-the-shelf or customize* - from OpenAI, Microsoft, etc.? Sage's AI team discovered the problem of off-the-shelf gen AI for finance firsthand. During his day two Sage Transform keynote, Harris explained:

When we first started [on gen AI] two years ago, I really challenged the AI team, 'Why are we not using something off the shelf? It can't possibly be that we can build this better than the Amazons or Microsofts of the world.' The answer they gave me was pretty straightforward: 'Those aren't designed for accounting workflows.' It's not only that they're

not accurate enough for accounting. It's also that they're not getting a lot of information to be for accounting.They don't understand the context around the data on the invoice; they often don't know what to do with it.

So that's how the Sage approach to finance models was sparked: take the big external foundation models, and combine them with Sage's own finance-specific models, to achieve the context finance users need. Harris continued:

So we do start with those foundation models. But then what our team does is they build all the follow-on models to solve the tricky problems for accounting. We've got 24 models now that we've deployed just for coding invoices. We've got a couple of models that do nothing but identify who the vendor is. We've got models for understanding the structure of the line items. Our models find hard-to-find values. Believe it or not, 'total' is really hard to find on an invoice off-the-shelf. OCR is like 25 or 30 percent on that one. We've also had to deal with taxes and currencies outside of the US across the world, and so we identify that.

Addressing the drawbacks of generalized LLMs for finance

During our post-keynote sitdown, I challenged Harris on why Sage would utilize a GPT-type model in any way for finance - especially given Sage's rigor around AI trust. As I said to Harris: *my eyebrows raise a little bit when I think about Sage using GPT-type models, because I think one of the big problems with trust is the training data itself. So I was really struck by the announcement that you made today about the world's first domain-specific accounting and compliance language model. What I was going to ask you today was: why not get around some of that by working with finance-specific Large Language Models that are not trained on Reddit, and are not trained on a variety of places that are going to have more wacky information that isn't really relevant to your purposes anyway?*

Harris responded that the foundation model choice is flexible. The key is understanding their limitations, and applying finance-specific models to refine their output:

So this won't necessarily be built on GPT, or any existing model. We might choose to work from an open source model. But part of the process is that additional layer of training, to optimize for the good stuff. One of the ways we build AI is to make sure that Large Language Models aren't doing things they shouldn't do.

So you never want a Large Language Model to do math, for example. If there's a need for a 100% accurate answer, you work to build systems, such that the Large Language Model can go to a skill that we've developed for it that's built specifically to handle that type of inquiry or perform that kind of task. This is the real challenge, actually, in bringing Copilot to market, is the AI can do amazing things; it can do absolutely amazing things. What's

required to bring it to market is developing all these skills, and ensuring that it operates in a safe and competent way.

Making customer-specific models a reality - how does the 'AI Factory' fit in?

But there is another component to Sage's AI strategy - customer-specific models. As Harris told attendees:

The other big part of this is that we're able to design models that are just for your business. So some of these models learn from the overall activity of the network. But when it comes to coding that invoice, we need a model that learns from your company - from the way you're set up your chart of accounts, your Dimensions, and the way your team works.

So we have seven models that we deploy just for your team that do the tricky bits around coding, which means we've got tens of thousands of models now in production, most of which are designed for individual customers, [which are retrained on a daily basis]. So the result of all of that is we now get 100% accuracy on the invoices we process, for about 68% of the invoices that flow through the AI system.

Harris credits the ability to deploy custom models to a framework he has informally dubbed the AI Factory. So what is this mysterious invention? Start with this:

In the AI world, software development is completely different.

Yes, Harris told me - that means AI software development is also completely different from the SaaS multi-tenant model Harris cut his teeth on with Intacct:

The challenge that I gave the team with our first real scale AI capability was 'Okay, we're now going from weekly to continuous. And we're going from every customer getting the same version of the code to every customer getting their own AI, and machine learning training infrastructure.'

Because there's some AI that learns best from lots of people. But a lot of AI just doesn't work that way; it needs to be trained on the unique characteristics and configuration and dimensionality in a business.

If you want to deploy AI this way, you build it from the ground up:

We took the same approach with AI as we did with the early days of Intacct: we built out the infrastructure that would enable us to do this in a scaled and automated way - deploy continuously trained and operated machine learning models into the [Sage Network], where they can learn from the global scale of the network, but also deploy into individual

companies, where they learn from those companies. With what I call the Sage AI Factory, we've got this way now to wrap it all into one secure container.

This is now Sage's AI development environment:

Every developer at Sage that's building AI capabilities, there's this foundation that enables them to automatically have continuous training - all the safety mechanisms built-in, the ability to do it for all customers, or for models that have to learn individually from from each unique customer.

Harris sees Sage's patented digital fingerprint as being instrumental here (this service provides a unique "fingerprint" for each vendor, based on Sage AI's understanding of their invoices). Harris:

The result of this is that for our customers, we're able to resolve about half of the mistakes that the off-the-shelf models make.

Harris mentioned that about 68 percent of the documents that move through Sage AI are processed with 100% accuracy. He thinks that can be improved:

We've created a continuous loop with our customers where their feedback, any changes that they make, any missing values that they provide, are fed back into our AI models so that they're consistently trained to become more accurate, as more customers provide more feedback over time.

My take - on hallucinations, SaaS, and making AI accessible to all customers

No surprise - Harris spoke of his team's concerted approach to prevent hallucinations: "*That's where the bulk of our primary research from our data science team is going.*" Sage tackles hallucinations in a couple ways: one is via what they call corrobation - essentially cross-checking LLMs via a "smoke test." That's done by comparing the LLM output with the data from the relevant workflow or transaction and flagging abnormalities. Yes, RAG (Retrieval Augmented Generation) is relied upon heavily here.

Harris also believes Sage can reduce hallucinations by simply not using LLMs in scenarios where they aren't needed, or are more likely to hallucinate. Thus the creation of so many finance-specific models inside of Sage.

But that brings us to the crux of it. *How is Sage able to deliver these AI services to all customers, when so many other vendors are focusing on AI for SaaS customers with structured/vanilla types of data environments?*

Harris says Sage can do this because accounting data has, by definition, a reasonable amount of structure. This data has likely been audited and is used for investor reports, etc. He points out that even Sage Intacct customers have quite a bit of variation due to the flexibility of the "Dimensions" they use to customize. So, in a sense, Sage had to solve this AI riddle for all their customers, including those on Intacct.

In terms of making AI innovations accessible, I think Harris underestimates what his team has accomplished here. It's refreshing to hear from vendors that will provide AI to all customers, rather than requiring core release upgrades and cloud migrations - though Harris does acknowledge that "*we may not always build out to the custom bits.*"

But Harris is certainly aware that bringing thousands of custom models to customers is a massive shift:

It's a huge part of why we can do this. The AI Factory is not a product. I am describing what we built internally to enable this.

That's why Harris says Sage is making not one big bet, but two: one on AI, and one on the Sage Network.

Harris is right: the ability to continuously retrain custom models is a major shift in how we think about modern enterprise software. Intacct is well-known for hammering at the advantage of a multi-tenant architecture versus pseudo-SaaS competitors. That hasn't changed - but a well-designed AI architecture can play by different rules. Unlike SaaS, you can have custom AI models and still retrain/update them easily - and Harris isn't the first AI expert who has been adamant on this point with me. "Custom" is typically a dirty word in a SaaS context, but AI is a different story.

Harris also gets credit for saying what so many vendors haven't: LLMs struggle with math, and struggle with finance out of the box. Even with Sage's refinements, 100 percent accuracy in all cases isn't currently possible. But Harris has plenty of examples where improved accuracy makes a major difference. That's why industry-tailored AI, with proper amounts of human oversight, is so much more intriguing than pushing enterprise questions into a GPT-type chat and hoping for the best.

The Sage Intacct customers I spoke with like Sage's AI emphasis and direction, with this caveat: finance problems are still human problems. Sage's talk of eliminating the monthly close electrified the audience - and provoked debate on whether this is really a realistic goal:

On X, I was challenged on whether the "continuous close" is achievable. But to me, the push for continuous close is not really about whether it's achievable anytime soon. It's about customers taking process/automation improvements seriously as an ambition. It's about moving closer to real-time views of finance. When customers do that, good things can happen - including cost savings on more efficient closes/consolidations:

Most aspects of this "continuous close" session were not about AI, or even brute force automation. It was about project discipline, underutilized features, and so-called "best practices." AI isn't magic - it's part of the bigger picture of how finance leaders need to change, and there is no time to spare. During our on-site interview, Sage CEO Steve Hare emphasized this point:

I think the most important thing is providing real solutions for real businesses... There's a lot of noise around AI; generative AI in particular has a huge amount of potential. But small/midsize businesses, business owners and their teams, they don't buy AI. They buy applications that help them solve pain points or give them business insights, etc.

That's really what I want people to take away. That's what we're focused on. Yes, of course, we're investing in the latest technology, but it's what we can get that technology to do for you to save you time - and give you business insights to help you run your business more successfully.

When it comes to modernizing the customer base with AI and more, Sage still has a lot to prove - but they've set an ambitious storyline to follow.

End note: for more on Sage's AI approach, check Aaron Harris' popular diginomica piece, Are we ready for the profound impact of generative AI on accounting and finance? Here's what lies ahead. I updated the above article Friday, March 29, 6pm US PT, with a number of small tweaks for reading clarity.

Image credit -

Photo of Aaron Harris at Sage Transform 2024 by David Gersten (@dsgersten) on X.