# Building Vertical AI

## Essential insights for AI startups ambitious enough to reimagine technology across vertical industries.

A book by
Bessemer Venture Partners

# What's inside

# Building Vertical AI: An early stage playbook for founders

Large language models (LLMs) have unlocked something traditional vertical SaaS never could: the ability to automate high-cost, language-intensive tasks that represent the vast majority of work in professional services. Lawyers reviewing discovery documents. Doctors writing clinical notes. Accountants preparing tax filings. These workflows were always too complex and nuanced for traditional software to handle. But LLMs have radically expanded what's possible.

The numbers paint a clear picture. Business and professional services, which are predicated on exactly these kinds of repetitive language tasks, account for 13% of the US's GDP. That's roughly 10x the size of the software market. Vertical AI isn't competing for IT budgets; it's competing for labor budgets. Unlike vertical SaaS, which typically captures a fraction of Fortune 500 IT spend, Vertical AI taps directly into the labor line of a P&L. That's why we believe Vertical AI represents a fundamentally larger opportunity than vertical SaaS ever did.

The momentum is undeniable. Growing giants such as Abridge (transforming clinical documentation), EliseAI (AI for housing and healthcare), EvenUp (automating personal injury law), and Fieldguide (reimagining audit workflows), Legora (AI for corporate law), and MaintainX (digitizes maintenance operations for manufacturing) are demonstrating what's possible when AI serves previously untouchable workflows.

In a previous report, we predicted that Vertical AI companies would hit $100M+ ARR within historic timeframes. We're already seeing this level of momentum from many AI supernovas, and we anticipate the first vertical AI IPO to likely happen within the next few years.

**This early stage playbook offers founders principles and frameworks for building defensible vertical AI products. We'll also offer guidance for teams on how to select the right business model, create competitive moats, and prove ROI quickly in your early stage journey of transforming language-intensive workflows in underserved industries.**

## Getting started on building your Vertical AI product

The best vertical AI products don't start with a technology thesis—they start by solving a specific workflow problem in an industry desperate for better tools. Some of the greatest SaaS companies followed this pattern. Shopify, among many others, all started as internal solutions before spinning out to become category leaders.

Many startups are built by industry insiders who have deep experience in their vertical, but we have also witnessed the power of the outsider, which includes product and technological experts that solve entrenched industry challenges with GenAI and a fresh approach.

For example, Shivdev Rao, M.D. started his career as a cardiologist before founding the healthcare AI company Abridge and Jin Chang was an auditor before founding Fieldguide, the advisory and audit automation platform. And then there's Max Junestrand of Legora who isn't a lawyer, and yet has built the AI solution helping law firms collaborate better. Minna Song of EliseAI follows a similar profile; she saw the potential to build AI agents to improve the efficiency of healthcare and housing, despite never working in property management.

A Vertical AI founder's "edge" doesn't always come from industry experience on the resume, rather it's the systems thinking, imagination, and deep customer empathy to find a vertical workflow or task that was previously not automatable before GenAI.

# Four frameworks to help you select your product idea

When honing in on a product idea in Vertical AI, make sure to think through the following four considerations.

**1. Clear ROI and technical feasibility**

Assess potential product ideas along two dimensions: business impact and technical feasibility. The sweet spot is high-impact workflows that AI can reliably execute.

Look for:

- Precisely-defined, repeatable processes your target customers have mastered. Newer or ambiguous processes should remain human-driven until standardized.
- Reasonable expectation AI can execute safely without degrading customer experience or creating compliance or security risks.
- Even better: workflows where AI introduces superhuman capabilities, such as analyzing 1000x more data, operating 24/7, detecting patterns invisible to humans.

Avoid:

- Processes requiring constant judgment calls that haven't been codified.
- High-risk workflows where mistakes compromise customer relationships or regulatory compliance.
- "Mundane but easy" tasks that don't create meaningful value.

**2. Insider vs. outsider advantage**

Are you solving a problem you've lived through, or addressing an industry as an informed outsider?

Insiders move faster initially—they understand workflows intimately, speak the customer's language, and have immediate credibility. The risk: conformist thinking that blinds them to disruptive approaches.

Outsiders face steep learning curves that burn months and capital understanding nuances insiders grasp intuitively. Yet outsiders can win by bringing specific expertise in applying emerging technologies—if they marry that knowledge with deep empathy for customer problems.

The failure mode is predictable: showing up with technology and assuming it'll work without obsessive attention to workflow integration. Successful outsiders don't just solve the problem technically; they understand how their solution fits seamlessly into daily operations, regulatory requirements, and existing systems.

Vertical AI typically rewards insider expertise more than horizontal SaaS did—you're reimagining complex, nuanced workflows in regulated industries, not just digitizing generic processes. That said, the best approach often combines both outsider creativity with insider networks for rapid validation. As we shared above, breakout Vertical AI leaders come from all backgrounds.

Here's what matters most: Intimate understanding of your industry's challenges, clarity on which workflows drive real ROI for customers, and a specific initial wedge you can dominate.

**3. Find your "magical feature" and earn the right to expand**

The best vertical AI products don't just automate existing workflows—they demonstrate a miracle-like advancement in how work gets done today. Your entry point matters, but what matters more is using that initial breakthrough to earn the right to expand into increasingly core workflows.
There are three strategic entry points, each with different risks and requirements:

**a. Tertiary workflows (the safer wedge)**

Starting with ancillary tasks—legal research, clinical documentation, construction estimates—offers clear advantages: less resistance to change, faster sales cycles, and obvious ROI that frees up capacity for higher-value work. Abridge wins because doctors desperately want administrative burden eliminated, creating an easy entry point.

But here's the critical insight: if you stay at the periphery too long, you become vulnerable to disintermediation. You must use that initial "magical experience" to build a right to expand into core workflows. Show such dramatic improvement in research or documentation that customers trust you to tackle brief writing or diagnostic support. The tertiary entry is a wedge, not a destination.

**b. Adjacent-to-core features (the balanced approach)**

If your breakthrough sits closer to core workflows—automating demand package generation for lawyers, underwriting analysis for insurers—you can still sell it as a standalone tool, but it needs sufficient value that users will tolerate a new solution in their critical path. Sixfold succeeds here by embedding directly into systems, making the integration seamless rather than disruptive.

The advantage is that you're already proximate to the core workflow, making expansion more natural. The challenge is that you need stronger proof of reliability since mistakes impact business-critical operations. Early customers require more convincing, but once proven, your moat is deeper.

**c. Core workflows (the Toast playbook)**

Going after the "heart/lung machine"—the absolute center of how customers operate—can work, but the miracle must be undeniable. Toast didn't improve restaurant payment processing incrementally; they reimagined the entire point-of-sale system when competitors were still using outdated hardware.

This approach demands missionary zeal from your team. Your first customers take a genuine leap of faith, and they'll only do that if you demonstrate you're on a mission to transform their lives, not just sell software. The Toast team sleeping at customer restaurants, committing code on laptops behind the counter —that level of commitment convinced early adopters to bet their operations on an unproven solution.

Realistically, most founders should start with tertiary or adjacent-to-core "miracles" and expand from there. But regardless of entry point, the principle holds: demonstrate something miraculous, then use that credibility to expand before competitors catch up.

Keep in mind, speed matters more than ever compared to SaaS.

Traditional startup wisdom says "stay focused on one product until fully established." That made sense when each new feature required months of development. AI changes the equation. Once your initial miracle proves reliable, immediately begin planning adjacent expansions—your "Second Act."

This layer-cake approach (à la Procore, Toast, ServiceTitan) builds switching costs and deepens integration before competitors respond. Model capabilities evolve weekly, and the window to establish category leadership is measured in quarters, not years. Use your initial product to earn the right to expand, then move fast.

**4. Progressive delegation vs. complete replacement**

AI doesn't need to fully replace a workflow to be valuable. In fact, making complete automation the goal can be detrimental—it creates unrealistic expectations that kill momentum when reality falls short.

Instead, aim for "progressive delegation." Start by automating a manageable chunk of the workflow (the most time-intensive, lowest-value slice.) Let humans handle the rest initially. As you learn from production usage, gradually automate additional steps.

**This approach has multiple benefits:**

- Ships faster with tighter feedback loops
- Reduces risk by keeping humans in the loop for high-stakes decisions
- Allows customers to standardize remaining manual steps, making them automation-ready later
- Creates expansion revenue opportunities as you automate more of the workflow

Here's an example: An AI audit tool might start by automating evidence gathering and preliminary analysis, while auditors handle final judgment calls and client communication. Over time, as the system proves reliable, it can take on more of the review process—but the initial product delivers immediate value without requiring perfect automation.

## Selecting high-impact use cases: The three-part test

Before committing to a product idea, ensure it passes three tests:

**1. Does your product idea provide enablement ROI?**

Does your automation unlock entirely new capabilities customers couldn't do before? This is more valuable than pure productivity gains.

- Enablement: Custom demos without engineering time, analysis of datasets too large for humans, coaching based on 1000x more conversations
- Cost savings: Offset future hiring, prevent expensive turnover (e.g. Abridge saves millions in physician retention costs)
- Productivity gains: Save time on tasks (only valuable if reinvested into strategic work)

Enablement ROI often matters most for enterprise sales—it's a capability unlock, not just an efficiency play.

**2. Does your product target processes that are standardized and repeatable?**

Look for processes your target customers execute regularly, the same way every time, with consistently positive results. If the process requires constant improvisation or judgment calls that haven't been codified, it's not ready for automation.

Test this by documenting the workflow in extreme detail. If you find yourself writing "it depends on..." frequently, the process needs more standardization before AI can handle it reliably.

**3. Does the use case have narrow scope and low risk (initially)?**

Start with automations where mistakes are recoverable and don't compromise customer relationships, privacy, security, fairness, or compliance. Even if you believe risks are low, wait until you have production experience before tackling high-stakes workflows. Once you've proven reliability with lower-risk use cases, expand to higher-value (and higher-risk) workflows with confidence and customer trust already established.

## 10 principles for Vertical AI founders to follow

We see business defensibility comes from understanding workflows deeply enough to automate them reliably, integrating tightly into existing systems, and pricing for the value you create. (Revisit Part IV of our series for a deeper dive into each principle.)

1. Customer-centric automation: Build solutions only where automation aligns with customer needs and context, not just possibility.
2. Avoid commoditized features: Focus on differentiated, integrated workflows rather than features that competitors can easily replicate.
3. Leverage AI for superhuman tasks: Identify and implement AI in areas where it can operate at scales or speeds unattainable by humans.
4. Quantifiable ROI unlocks value: Demonstrate clear revenue gains or cost reductions to drive adoption and loyalty.
5. Innovate on business models: Embrace new delivery methods and pricing enabled by AI automation to access broader market segments and expand margins.
6. Target niche and underserved markets: Initial competitive advantage often lies in overlooked, high-ROI areas.
7. Customize for nuanced requirements: Serve complex buyer needs such as compliance or security to erect defensible barriers.
8. Technical moat comes from multimodality: Competitive edge increasingly depends on combining data types and workflow integrations, not proprietary models alone.
9. Build modular and adaptable systems: Ensure tech infrastructure can flexibly incorporate the best models as AI evolves.
10. Prioritize data quality over quantity: Early success depends on high-quality, relevant data, which compounds in value as the business scales

## A call for the next generation of Vertical AI founders

Tomorrow's titans are being built today—but countless others are still ideas waiting to be imagined. If you're an aspiring founder working in an industry in desperate need of revitalization, start by automating your own workflows in your everyday work. If you can make solutions work for yourself, you've validated your first product.

If you're already working on a Vertical AI application, we would love to hear from you. Email our team at VerticalAI@bvp.com.

PART I

# The future of AI is vertical

We explore how a new class of LLM-native applications are unlocking markets previously out of bounds for legacy SaaS — and the massive value creation already underway.



After building one of the largest vertical SaaS portfolios in venture, we learned that great vertical software companies can unseat incumbents, transform industries and the way people work, and become highly profitable, generational businesses. Now that the top 20 public vertical SaaS companies in the US have a combined market capitalization of ~$300 billion, this perspective feels obvious, but when Bessemer started investing in Mindbody, Shopify, Procore, and others back in the early 2010s, vertical software startups were seen as "sleepy" and their potential was uncertain.

The meteoric rise of these businesses over the past 15+ years and the advancements in AI during that same period has set the stage for an exciting new development in the vertical software landscape: Vertical AI. This all came to a head in 2023, when we saw a new class of LLM-native applications harnessing novel business models and AI capabilities in order to serve functions and even entire industries that didn't meaningfully benefit from the previous wave of vertical software.

Unlike their predecessors, these vertical AI applications are able to target the high cost repetitive language-based tasks that dominate numerous verticals and large sectors of the economy — such as legal, healthcare, and finance — that were largely out of bounds for legacy vertical software. Given Vertical AI's ability to both capture new markets and tap into more sizable TAMs within those markets, we predict that Vertical AI represents an even larger market opportunity than that of legacy vertical SaaS.

There are already AI-first teams beginning to solve industry-specific problems using LLMs and generative AI. And then there are vertical SaaS leaders continuing to serve businesses with software solutions. (For the latter, it's time to consider incorporating AI into your product, if it's not been incorporated already. Look to Intercom, Zapier, and Canva for inspiration.)

In this first installment of our Vertical AI series, we focus on the dynamics driving this promising, fast-moving, and highly competitive category.

Vertical AI is changing startup physics in the enterprise software landscape; we're seeing this emergent category address an even more massive TAM, offer new capabilities, serve new sectors of the economy, and grow at unprecedented rates. In Part I, we distill why this moment is so significant in software history, the key opportunities driving this category's massive value creation, and the core and supporting workflows underpinning these emerging businesses, as well as highlight a few promising startups already gaining speed within their industries.

**Vertical AI is changing startup physics in the enterprise software landscape; we're seeing this emergent category address an even more massive TAM, offer new capabilities, serve new sectors of the economy, and grow at unprecedented rates. In Part I, we distill why this moment is so significant in software history, the key opportunities driving this category's massive value creation, and the core and supporting workflows underpinning these emerging businesses, as well as highlight a few promising startups already gaining speed within their industries.**

## Why build now — the potential of Vertical AI

You might be thinking, "A sea change is certainly coming, but it's not here yet and we still have time." While it's true that horizontal SaaS preceded vertical SaaS by a significant margin in the previous SaaS wave (with Salesforce going public an entire decade before Veeva Systems), we also know that many frameworks and timelines for SaaS businesses don't always map perfectly onto AI.



We're excited by the early data we're seeing on Vertical AI business models. We'll note that for the most part Vertical AI players are leading with functionality that isn't competing with legacy SaaS. The utility of these applications is typically complementary to a legacy SaaS product (if one exists at all) and thus doesn't have to replicate and displace an incumbent.

Analyses of our Vertical AI portfolio hint at the strength of this new class of applications. LLM-native companies in this cohort (with founding dates of 2019 to present) have quickly reached 80% of the average contract value (ACV) of the traditional core vertical SaaS systems, and are growing ~400% year-over-year, while still maintaining a healthy ~65% gross margin.

**Based on the growth rates of more mature Vertical AI startups in this category, we predict we'll see at least five Vertical AI companies with $100M+ ARR within the next two to three years. We also anticipate the first Vertical AI IPO in the next three years.**

We're already seeing exit activity among vertical AI companies through M&As. In 2023, Thomson Reuters acquired CaseText for $650M and a year later, DocuSign acquired Lexion for $165M. Incumbents are building as well as buying. In a recent survey, a majority of Bessemer investors who collectively oversee Bessemer's vertical SaaS portfolio reported that companies across regions and sectors are rapidly incorporating AI features into their products.

## Key opportunities in Vertical AI

There are numerous reasons why we see so much potential in Vertical AI and believe that current conditions are ripe for success similar to (and likely exceeding) what we saw in the previous waves of vertical software. Three opportunities stand out:

### Expand total addressable markets (TAMs)

**Historically, software builders have focused on the largest TAMs, ignoring niche categories where it would be difficult to build a big business. Vertical AI is able to unlock markets that previously would have been perceived as too small to build a sustainable SaaS business because they are significantly increasing the scope of value delivered through AI.**

**Take EvenUp, which automates demand letter generation for personal injury attorneys. By allowing firms to take on more customers at a lower cost — and thereby increasing margins — EvenUp has exceeded the traditional SaaS TAM that you'd expect of a solution that improves workflows for demand letter management.**

**The US Bureau of Labor Statistics cites software spend as 1% of the US GDP, and the Business and Professional Services industry — dominated by repetitive language tasks — at 13%. Based on this telling statistic as well as our own surveys, research, and observations, we predict that Vertical AI's market capitalization will be at least 10x the size of legacy Vertical SaaS as Vertical AI takes on the services economy and unleashes new business models uniquely capable to serve this category. (We'll share more on those new Vertical AI business types later on in our series.)**

### Unlock new functions and verticals

Vertical software replaced outdated and cumbersome systems and brought many industries — from construction to hospitality — online for the first time. But these developments left large swaths of



the economy behind. In many cases, the ROI of a software solution alone wasn't high enough to convince technophobic decision-makers and justify the upfront costs of setting up the required infrastructure, implementing new processes, and training employees.

Vertical AI startups have been gaining traction in markets that vertical software couldn't access by providing solutions that can radically improve workflows and often take over tasks entirely. As a result, we're seeing large incumbents within these industries become receptive to this technology, and sometimes even actively seek out AI-enabled tools out of concern that the competition will overtake them by adopting these tools first.

For example, in healthcare, an industry with notoriously long deal cycles for SaaS, providers are adopting solutions such as Abridge — which turns patient-doctor conversations into clinical notes — and ClinicalKey AI — an AI-powered medical search platform —- to take over busywork and support clinical decision-making. Law firms, which rarely even use CRMs, have also already begun adopting co-pilot based solutions for contracting, demand summary generation, case intake, and other time-intensive tasks.

### Provide unprecedented value

There's a future where, depending on capabilities, AI applications could be integrated into every industry in the economy, from home services through accounting. That said, the potential penetration of AI will vary by industry.

The most attractive markets to build Vertical AI companies will likely be in contexts where AI facilitates or completes work that was previously impossible or too expensive to achieve with human labor alone, and so the work was not being done or being done poorly. A common, successful use case for AI is automating or streamlining workflows by reviewing significantly more data than humans would previously have been able to audit.

For example, Axion Ray helps manufacturers by analyzing large volumes of product data across IoT & telematics, field failures, production, and supplier data. Similarly, JusticeText automatically reviews hundreds of hours of camera footage to help public defenders build their cases — something that's extremely time-intensive for lawyers to do during discovery and which also takes away focus from building cases. Later on in this series, we'll explain how the rise of multimodal models is allowing Vertical AI businesses such as JusticeText to go beyond purely text-and-data-workflows and also leverage voice, video, images, and other sources of input.

## Vertical AI for core vs. supporting workflows

Winners in the previous wave of vertical SaaS created cloud platforms that were purpose-built for underserved markets, with many adding more and more integrated products and services over time to eventually produce a "layer cake" that provides an all-in-one solution for a given vertical and drives continuous growth (as Procore did for construction and Toast did for restaurants, for example).

As we've discussed in this article, vertical AI businesses can access larger TAMs within a given market by offering high-ROI solutions and therefore don't always need to have as large of a product scope as their predecessors did in order to build successful businesses. In fact, some promising vertical AI startups are able to break into industries and drive returns by addressing just one or two of the target customer's workflows.



We divide these workflows into two categories: core and supporting. Core workflows are those that are a primary function of a job; for example, financial modeling for an investment banker or contract drafting for a lawyer. Supporting workflows are those incidental to a job or business but still necessary; for example, marketing and patient relationship management for a dentist (e.i. Weave) or freight procurement for a shipper (i.e. GoodShip).

As companies of all types digitize in every industry, they need a software stack that helps them not only do the job, but also run and operate the business. Now, in the AI era, there are opportunities for companies in every industry to leverage AI in their trade or service and their business operations.

Lastly, building on recent progress in ML and hardware, a new wave of AI-native and AI-embedded startups is emerging—these companies either leverage AI/ML from the ground up or use it to augment their existing capabilities. Unfortunately, much of current data infrastructure and tooling is still not optimized for AI use cases. Similar to forcing a square peg into a round hole, AI engineers have had to create workarounds or hacks within their current infrastructure.

## Core workflows

Today, text-and quantitative-heavy work have the highest propensity for automation. That's why we tend to see more vertical AI solutions addressing the core workflows of industries that are dominated by traditional office work —- such as in legal and professional services — rather than those that require significant manual labor — such as home services and manufacturing. For example, portfolio company Fieldguide is revolutionizing the core workflows of auditors involved in diverse projects such as SOC 2 engagements, financial audits, PCI DSS assessments, and internal audits. By leveraging automation and generative AI, Fieldguide enhances auditor efficiency, leading to significant productivity gains in tasks where professionals spend most of their time.

However, just because a core workflow has a high propensity for automation doesn't necessarily mean it's a good use case for AI. A target customer's desire to automate a given workflow matters too, and that will vary significantly across sectors. For instance, investment bankers may use AI to automate the tedious process of slide creation, but be unlikely to use a voice AI that gives presentations to clients because of the importance of relationships in the space.

## Supporting workflows

Supporting workflows may be better targets for Vertical AI specifically because they are ancillary to the job of the target customer, and therefore are typically the type of work that can be delegated to and completed satisfactorily by AI. For example, a doctor has both expertise and interest in treating patients but less in, say, notetaking and paperwork (increasing doctors' "pajama time") or even ordering medical supplies. That's likely why we see a high market demand for AI solutions addressing supporting workflows across back office operations, sales, procurement, finance, and other functions.

However, addressing supporting workflows with AI is not without challenges. For one, many tech-forward horizontal incumbents in these sectors have already begun incorporating AI into their platforms, and vertical AI startups will need to deliver meaningfully better solutions in order to compete.

On the upside, vertical-specific AI startups may be better poised to understand sector-specific nuances and integrate with underlying systems (such as CRMs) and therefore be able to create an experience that's harder for horizontal competitors to replicate with a general LLM.

For example, an AI solution built specifically for home services can identify a customer's problem and route a technician to fix a solar panel faster and more effectively than a horizontal solution that can only make an appointment at the customer's request. Still, founders will need to pay close attention to the potential TAM of any solution given that these forms of defensibility may come at the expense of market size and therefore require the layer cake approach mentioned above.

Whether building AI for core or supporting workflows, founders need to have good judgment, a deep understanding of customer needs, effective feedback channels, and a clear grasp of the regulatory landscape in order to identify the specific sectors and tasks that are well-suited for an AI solution. Remember: just because something can be automated, doesn't mean it should.

## A note on Vertical AI moats

For example, an AI solution built specifically for home services can identify a customer's problem and route a technician to fix a solar panel faster and more effectively than a horizontal solution that can only make an appointment at the customer's request. Still, founders will need to pay close attention to the potential TAM of any solution given that these forms of defensibility may come at the expense of market size and therefore require the layer cake approach mentioned above.

Whether building AI for core or supporting workflows, founders need to have good judgment, a deep understanding of customer needs, effective feedback channels, and a clear grasp of the regulatory landscape in order to identify the specific sectors and tasks that are well-suited for an AI solution. Remember: just because something can be automated, doesn't mean it should.

## Up next: Multimodal innovation

Vertical AI excites investors and entrepreneurs, but it also terrifies them. With jaw dropping products, early breakout growth, SaaS margins, and healthy TAMs come an unprecedented pace of development, intense competition, elevated valuations, and risks from incumbents who are not asleep at the wheel — and that's just in the first inning. Given these competitive dynamics, we're excited about the future of multimodal AI. In the next installemnt, we cover developments in multimodal model architecture, exciting multimodal voice and vision applications, and the promise of AI agents.

PART II

# Multimodal capabilities unlock new opportunities in Vertical AI

Vertical AI applications born out of novel audio, voice, and vision capabilities will fundamentally change the way we work.

Since ChatGPT first came onto the scene and captivated entrepreneurs (and the world), we've witnessed a massive explosion in products and services using LLMs to address the "lowest hanging fruit" use cases for generative AI: text-based tasks — spanning everything from creating legal contracts and job descriptions to drafting emails and website copy.

Demand for text-based AI solutions remains high. AI can take over time-consuming tasks like creating first drafts, thereby refocusing employee efforts to more complex functions. But much more of our day-to-day work requires data types and capabilities other than just text, such as speaking to customers and reasoning over complex images and graphical data. Today, use cases like these are no longer off the table.

**The emergence of multimodal models has created opportunities for vertical AI to impact a much larger share of the economy than previously imagined by expanding beyond text-based tasks and workflows. In Part II, we report on new models that support a variety of data types across audio, video, voice, and vision, promising early applications of new and improved voice and vision capabilities, and the potential of AI agents to change how businesses operate.**

## Exciting developments in multimodal architecture

In the past 12 months, new models have emerged that demonstrate significant advancements in terms of their ability to understand context and reduce hallucinations, as well as their overall reasoning capabilities. The performance we're seeing across speech recognition, image processing, and voice generation in certain models is approaching (or, in some cases, surpassing) human capabilities, unlocking many new use cases for AI.

## Voice capabilities

We've seen rapid progress made on two core components of the conversational voice stack: speech-to-text models (automatic speech recognition) and text-to-speech models (generative voice). Dozens of vendors are now providing models with these capabilities, which has led to a flurry of new AI applications, particularly in the case of conversational voice.

Most of these applications rely on what's called a "cascading architecture," where voice is first transcribed to text, then that text is fed into an LLM to generate a response, and finally the text output is fed back into the generative voice model to produce an audio response. Up until very recently, this has been the best way to build conversational voice applications. However, the approach has a few drawbacks — primarily that it introduces additional latency and some of the non-textual context (i.e., the end user's emotion and sentiment) gets lost in the transcription process.

As of the time of this writing, a new generation of speech-native models are being released including OpenAI's Realtime API, which supports speech-to-speech interactions via GPT-4o, as well as several open-source projects such as Kyutai's Moshi. Developing models capable of processing and reasoning on raw audio has been an active area of research for many years and it's been widely acknowledged that speech-native models would eventually replace cascading architecture.

Speech-native models have substantially lower latency (< 500 milliseconds) than previous models. They can also capture much more context from users (i.e., their tone, sentiment, emotion, etc.) and generate responses reflecting that context, making exchanges feel more natural and increasing the likelihood that they address the user's needs.

Over the next few years, we anticipate a step function change in the speed and quality of conversational voice applications, as more of them are built on these new and improved models.

## USE CASES FOR VOICE

There's already been dramatic progress in transcription-based applications, where speech-to-text models are a bit more mature. As a result, there's been notable early progress in end-to-end conversational voice agents — something we consider the "next frontier" for voice AI solutions.
**Let's look at four initial use cases:**

### 1. Transcription frees up time for users to facilitate next steps in workflows

Transcription frees up time for users to facilitate next steps in workflows: Bessemer portfolio company Abridge has pioneered a best-in-class medical transcription application that can generate medical notes based on clinical conversations, and identify appropriate follow-ups including prescription ordering, scheduling appointments with specialists, and referencing billing codes. When doctors don't have to complete those tasks manually, they can redirect that time and attention towards patient care.

Another great example is Rillavoice, a company that's bringing AI to the home services vertical. Rillavoice's transcription application records conversations between salespeople and customers for training purposes so that sales managers can still provide valuable coaching feedback without having to go on very time-consuming in-person "ride-alongs."

### 2. Fielding inbound calls to capture incremental revenue

One most compelling use cases for end-to-end voice agents we've seen so far is inbound sales, particularly when the solution is purpose-built for a specific vertical (i.e., home services businesses or automotive dealerships). Voice agents can ensure that a business never misses a valuable lead by fielding customer calls after hours or when other sales representatives are busy. Some solutions are able to book an appointment for the customer and even interact with the customer's system of record to quote a price. These capabilities — combined with the dramatic improvement in conversational voice models compared to prior voice bots — have made it possible for some AI sales agents to close inbound leads at an impressive rate, without requiring interactions or interventions from a sales representative.

### 3. Upleveling customer success experience with AI

Customer support has long been a target of automation, but many users found prior versions of Interactive Voice Response (IVR) technology quite frustrating to use. Modern voice agents have proven to be much more effective. While traditional IVR products could only understand a customer's intent in response to specific phrasing, modern voice agents are able to provide a correct answer regardless of how customers ask questions or make requests. And as with all these use cases, automating phone calls gives time back to customer service representatives to focus on solving complex customer problems and answering nuanced questions (vs. FAQs).

### 4. Automating outbound calls to increase top of funnel

Multiple solutions have emerged to automate outbound calls for sales and recruiting teams. Typically, the voice agents use the customer's stated criteria to identify the highest-potential sales leads or candidates, make the initial call to the leads, and then route them to the next meeting with a salesperson or recruiter.



Having AI take over the outbound workflow significantly increases the number of leads that can be contacted and the company's top of funnel as a result. And with their time freed up, salespeople and recruiters have a better chance of closing the highest potential leads.

It will be critical to monitor regulation in this space that limits unwanted robocalls and to ensure that solutions facilitate outbound calls only to leads and candidates that opt in to sales outreach.

Across all voice use cases, we expect that low latency and understanding a user's sentiment and emotion will become table stakes, and that more sophisticated solutions will differentiate along other dimensions, such as orchestrating conversations across multiple underlying models in real-time to optimize cost and performance; supporting omnichannel communications, multiple languages, and real-time translation; and building effective conversational guardrails, particularly in highly-regulated use cases.

## Vision capabilities

In vision, we've seen the development of models like GPT-4 with vision (GPT-4V) that can interpret images and respond to questions about them, as well as multimodal models like GPT-4o that can process raw images and video. It's expected that GPT-5 will also be able to reason more accurately across an increased context window, deriving nuanced insights from image inputs and potentially adding video processing capabilities too. Google's multimodal model Gemini 1.5 Pro can already understand input across both image and video and retain contextual understanding up to one million input tokens.

We expect that these and similar models will continue to improve in performance and come down in cost — great news for application builders.

### USE CASES FOR VISION AND VIDEO

Initial use cases for vision within vertical applications tend to fall into one of four categories: data extraction, visual inspection, design, and video analytics. While data extraction is the most mature use case for vision models so far, we're seeing progress in other areas, and still just scratching the surface of all the other potential use cases where vision models could be applied.

#### 1. Data extraction from pictures, PDFs, or images of other unstructured documents

AI can relieve humans of tedious data entry tasks and open up downstream workflows by applying structure to currently unstructured data. For example, Raft's platform for the freight forwarding industry uses a combination of computer vision and LLMs to extract critical information from PDF invoices, populate its customers' enterprise resource planning platforms (ERPs), and automate downstream tasks like invoice reconciliation and preparation of customs declarations.

#### 2. Augmenting jobs that currently involve human visual inspection

A number of companies have emerged that use AI to help streamline manual visual inspection processes and deliver faster results.

The AI construction platform xBuild generates scope-of-work packages for residential construction and restoration projects, and then partners with insurance companies to get them approved for reimbursement. xBuild uses photos of the damaged roofs and blueprints of houses to generate reports that outlines the scope of repair that will be required to restore roofs to their proper condition, in accordance with local building codes. Other applications have used AI and computer vision to automate the process of quality assurance reviews in construction drawings, helping to catch errors early to prevent costly changes later down the line in the construction process.

### 3. Generating 2D and 3D designs

There has been a steep increase in AI platforms serving the architecture, engineering, and construction (AEC) industry. Some companies are using AI to create feasibility assessments that combine visual depictions of proposed sites (buildings, parking lots, etc.) with the associated costs of supplies, adjusting the former with constraints of the latter, and vice versa. Other solutions like Snaptrude create detailed 3D designs of buildings, taking over the repetitive work typically done by structural engineers and giving them time back to focus on higher-level design work (rather than tedious tasks like putting pipes in the right place). Automating aspects of detailed product and infrastructure design not only saves customers valuable engineering time but can also strengthen sales proposals and increase project win rates.

### 4. Video analytics

Models that generate and/or understand video are the least mature among vision models, but they're improving rapidly. For example, video understanding models have become fairly capable when it comes to object tracking, classification, and even natural language search of video content. There have already been some compelling commercial applications built on top of these models, such as those that monitor video feeds for safety violations in manufacturing or industrial settings. But given how fast video models are maturing, we expect to see even more impressive applications in the coming years, and an expansion into more use cases, particularly in robotics where video understanding is a critical component of robotic perception.

Across all vision use cases, founders should avoid mistaking complexity for value. While solutions may be the most defensible when they automate particularly complex workflows—say, the creation of detailed 3D designs vs. 2D error checking—the value to customers will almost always be directly tied to how well the automation fits within a user's existing workflow.

If a design automation solution requires burdensome integrations with difficult-to-replace core systems (such as Revit) and has low initial ROI, it will be hard to drive sales and adoption, regardless of how robust the solution is. Early-stage companies may be better off starting with a product that's less technically complex and narrower in scope and then expanding from there. Of course, the best path will vary sector by sector and use case, but the trade-offs are important to keep in mind.

## USE CASES FOR VISION AND VIDEO

While early hype around AI agents has fallen short of reality, we've witnessed real progress recently, as teams find ways to more effectively constrain tasks for AI agents in order to reduce compounding errors in multi-step reasoning.

We're especially bullish on agents given the amount of research and resources being dedicated to reasoning-focused foundation models, such as OpenAI's o1. Most LLMs focus purely on predicting the next token based on patterns seen in training data, but models like o1 take a fundamentally different approach to problem-solving. These models are designed to "think" more at inference time, using chain-of-thought reasoning to better plan and assess their approach before arriving at an answer. While early, these models are demonstrating impressive performance on more complex reasoning tasks.

Today, agents are playing a valuable role in text, voice-based, and vision-based workflows that involve repetitive tasks and communication (as we show below). But over the coming year, we expect applications built on newer reasoning-based models to emerge and deliver on the true potential of AI agents: addressing complex workflows autonomously.

### 1. Data extraction from pictures, PDFs, or images of other unstructured documents

Many companies have come out with AI agents that can source and contact potential customers for sales teams. What's promising about these agents is that they're able to conduct a significant amount of research to identify high-quality prospects (through a detailed web search of the target company, its employees, and relevant industry news), and then use those research findings to craft relevant and highly personalized emails. Because agents can effectively execute the research and outreach portions of the job while maintaining relatively high quality, it allows sales representatives to redirect time towards closing warm leads.

### 2. Augmenting jobs that currently involve human visual inspection

A number of companies have emerged that use AI to help streamline manual visual inspection processes and deliver faster results. The AI construction platform xBuild generates scope-of-work packages for residential construction and restoration projects, and then partners with insurance companies to get them approved for reimbursement. xBuild uses photos of the damaged roofs and blueprints of houses to generate reports that outlines the scope of repair that will be required to restore roofs to their proper condition, in accordance with local building codes. Other applications have used AI and computer vision to automate the process of quality assurance reviews in construction drawings, helping to catch errors early to prevent costly changes later down the line in the construction process.

### 3. Generating 2D and 3D designs

There has been a steep increase in AI platforms serving the architecture, engineering, and construction (AEC) industry. Some companies are using AI to create feasibility assessments that combine visual depictions of proposed sites (buildings, parking lots, etc.) with the associated costs of supplies, adjusting the former with constraints of the latter, and vice versa. Other solutions like Snaptrude create detailed 3D designs of buildings, taking over the repetitive work typically done by structural engineers and giving them time back to focus on higher-level design work (rather than tedious tasks like putting pipes in the right place). Automating aspects of detailed product and infrastructure design not only saves customers valuable engineering time but can also strengthen sales proposals and increase project win rates.

### 4. Video analytics

Models that generate and/or understand video are the least mature among vision models, but they're improving rapidly. For example, video understanding models have become fairly capable when it comes to object tracking, classification, and even natural language search of video content. There have already been some compelling commercial applications built on top of these models, such as those that monitor video feeds for safety violations in manufacturing or industrial settings. But given how fast video models are maturing, we expect to see even more impressive applications in the coming years, and an expansion into more use cases, particularly in robotics where video understanding is a critical component of robotic perception.

Across all vision use cases, founders should avoid mistaking complexity for value. While solutions may be the most defensible when they automate particularly complex workflows—say, the creation of detailed 3D designs vs. 2D error checking—the value to customers will almost always be directly tied to how well the automation fits within a user's existing workflow.

If a design automation solution requires burdensome integrations with difficult-to-replace core systems (such as Revit) and has low initial ROI, it will be hard to drive sales and adoption, regardless of how robust the solution is. Early-stage companies may be better off starting with a product that's less technically complex and narrower in scope and then expanding from there. Of course, the best path will vary sector by sector and use case, but the trade-offs are important to keep in mind.

## The promise of AI agents

While early hype around AI agents has fallen short of reality, we've witnessed real progress recently, as teams find ways to more effectively constrain tasks for AI agents in order to reduce compounding errors in multi-step reasoning. We're especially bullish on agents given the amount of research and resources being dedicated to reasoning-focused foundation models, such as OpenAI's o1. Most LLMs focus purely on predicting the next token based on patterns seen in training data, but models like o1 take a fundamentally different approach to problem-solving. These models are designed to "think" more at inference time, using chain-of-thought reasoning to better plan and assess their approach before arriving at an answer. While early, these models are demonstrating impressive performance on more complex reasoning tasks.

Today, agents are playing a valuable role in text, voice-based, and vision-based workflows that involve repetitive tasks and communication (as we show below). But over the coming year, we expect applications built on newer reasoning-based models to emerge and deliver on the true potential of AI agents: addressing complex workflows autonomously.

### 1. Sales and marketing

Many companies have come out with AI agents that can source and contact potential customers for sales teams. What's promising about these agents is that they're able to conduct a significant amount of research to identify high-quality prospects (through a detailed web search of the target company, its employees, and relevant industry news), and then use those research findings to craft relevant and highly personalized emails. Because agents can effectively execute the research and outreach portions of the job while maintaining relatively high quality, it allows sales representatives to redirect time towards closing warm leads.



### 2. Negotiations

AI agents have shown promise in automating negotiations across multiple parties. Companies like Pactum have developed AI agents that can negotiate legal and commercial terms for supply chain use cases. Pactum's application maps out the value function for the user while the agent conducts simultaneous negotiations with suppliers to optimize deal terms. We've seen similar approaches taken by other vertical AI companies in the sales and promotions space. Here, agents negotiate with buyers and suppliers on set criteria such as discounts for bulk purchases or rapid payment plans.

### 3. Investigations

Enterprise cybersecurity teams are often overwhelmed by the high volume of security alerts they receive, but there are now AI agents that can assist with the initial phase of an alert investigation. That includes: gathering information about an event from multiple disparate systems, researching the malicious behaviors that might have been involved, and summarizing the incident and grading its severity. While most teams tend to use agents for lower-stakes workflows, it's clear that more sophisticated agents can (and likely will) address more and more workflows over time that require information gathering and synthesis.

We believe that agents addressing tasks and workflows that require more complex reasoning across multiple modalities will be significantly more defensible than solutions that don't. In particular, we're seeing that it's possible to drive higher performance in agentic workflows through clever architectural decisions and by stitching together the right models, guardrails, and feedback loops to deliver consistent outcomes. Agent performance is not purely based on the scale of data and compute that's thrown at a problem (as is the case with LLM training), and so this is a more compelling opportunity for early-stage startups. In all cases, it will be key to strike the right balance between building a technical moat and ensuring flexibility given the fast-paced development of the underlying models.

## Vertical AI expands its horizons

We believe that agents addressing tasks and workflows that require more complex reasoning across multiple modalities will be significantly more defensible than solutions that don't. In particular, we're seeing that it's possible to drive higher performance in agentic workflows through clever architectural decisions and by stitching together the right models, guardrails, and feedback loops to deliver consistent outcomes. Agent performance is not purely based on the scale of data and compute that's thrown at a problem (as is the case with LLM training), and so this is a more compelling opportunity for early-stage startups. In all cases, it will be key to strike the right balance between building a technical moat and ensuring flexibility given the fast-paced development of the underlying models.

## Up next: Novel business models

Advancements in LLMs and generative AI have driven business model innovation as much as product innovation. They've catalyzed novel software business models that have opened up opportunities in industries that were previously off limits for vertical software, as well as facilitating new use cases that have allowed existing vertical software incumbents to continue building a "layer-cake" of products and services. In the next article, we'll take a deep dive into three of these emerging business models—copilots, agents, and AI-enabled services—and the roles, potential applications, and pricing strategy for each model.

PART III

# Business model invention in the AI era

Copilots, agents, and AI-enabled services are three new models delivering the benefits of artificial intelligence to vertical industries.



Much like cloud computing transformed vertical industries over the past two decades, the human-like capabilities of LLMs — processing text, images, videos, voice, and code — are making it possible for Vertical AI to achieve what was previously unimaginable for sectors of the economy that only benefitted marginally from previous waves of software innovation. Vertical AI companies aren't just streamlining workflows; they're taking on vast, high-cost, language-heavy tasks that dominate industries like legal, healthcare, and professional services — sectors representing a 10x larger TAM than the software market itself.

The scale of opportunity is unmatched, but the efficacy and future success of Vertical AI hinges on the strength of its business models. Delivering your product in the right format for your specific vertical alongside a pricing model that captures the value you're creating for customers is critical for the short and long-term viability of your business.

**In this installment of our Vertical AI Roadmap, we dive into three new business models defining the AI era — Copilots, Agents, and AI-enabled Services — and explore the innovative use cases and functions vertical AI companies are addressing today.**

## Copilots

The first incarnation of AI-native business models that we saw emerge were copilots. Copilots are AI applications that sit side-by-side with users as they go through workflows. In this way, AI copilots supercharge employee productivity while keeping the human user in the center of the workflow. Because copilots aim to increase individual employees' efficiency and efficacy, copilots are generally priced like traditional cloud software on a per-seat basis tied to the company's headcount.

This pricing model has already shown real momentum with public companies like Microsoft, Google, and Salesforce which have all been able to benefit from healthy price increases of their core products by offering copilot add-ons. For example, Microsoft's Office365 generally costs $15 to $30 per license, but getting access to their copilot product costs an additional $30, doubling or even tripling the price per seat.

### COPILOT USE CASES BY MODALITY

**Code** GitHub Copilot was one of the first widely adopted copilot applications in the market. Developers use Copilot to complete code, generate sample code, and ask questions, leading to productivity increases (reportedly as much as 55%) and higher developer job satisfaction. Our portfolio company Supermaven, which recently joined forces with Cursor, is another example. Its one million token context window understands a developer's code base and offers the best contextual code completion at lower latency, turning users into 10x more effective developers.

**Text** Many text-based copilots assist or complete tasks for users within document-intensive workflows. For example, the AI copilot Harvey makes it easy for lawyers to quickly summarize contracts and get immediate answers to questions about specific content. Similarly, our portfolio company Sixfold AI helps insurance underwriters collate and synthesize information across sources to understand a prospective client's risk profile.

**Voice** As we discuss in part two of this series, significant, ongoing progress is being made on core components of the conversational voice stack. Advancements in speech-to-text models (automatic voice recognition), in particular, have supported a new generation of voice-based copilots addressing transcription use cases. These copilots "listen" to conversations, transcribe conversations, and often offer additional features, such as analytics or preparing and even helping users complete the next steps in a workflow.

**Image** There are also AI copilots that generate images to streamline design processes in a variety of industries but are particularly becoming more prevalent in the architecture, engineering, and construction (AEC) industry. Workpack AI and Togal AI are visual copilots for estimators in the pre-construction process that automatically perceive, measure, and label project specs.

## Agents

While copilots help employees do their work, AI agents fully automate workflows for specific functions with minimal human intervention required. In this way, agents are decoupling software and productivity from human headcount and transforming what businesses can achieve. Agent solutions are not only being built by vertical AI startups; incumbents are diving in headfirst, with Salesforce launching Agentforce, their employee and customer-facing support platform.

Pricing strategy for agents is still developing. Given that agents can substitute future incremental headcount and allow existing employees to work on higher value tasks, many are priced based on the solution's output relative to human workers, and agent ROI is framed in terms of the money saved on expanding headcount.

### INITIAL USE CASES FOR AI AGENTS

**Software sales** Relevance AI's sales development representative (SDR) agent, Bosh, automates the process of identifying, researching, and contacting leads, and scheduling meetings.

**Recruiting** LinkedIn recently announced the launch of its first AI agent Hiring Assistant, which takes on certain workflows typically done by recruiters, including sourcing candidates, turning notes into drafts of job descriptions, and more.

**Customer support** Slang's voice AI fields phone calls for restaurants, answering simple questions, making reservations, and connecting customers to employees as needed. Similarly, Assort Health's AI agent call center for healthcare schedules patients appointments, reducing wait times and dropped calls.

**Back office functions** Tennr automates document and referral processing, data entry, and other manual workflows related to healthcare administration.

## AI-enabled services

By using software to automate work, AI-enabled services companies have the potential to deliver cheaper, faster, and more consistent services to the market and take share from incumbent services companies. Pricing for these services generally anchors to existing legacy service provider pricing, but, in many cases, automation allows AI companies to undercut existing providers and still retain higher margins, thanks to AI's lower cost structure. For example, EvenUp charges per demand letters generated, which is less than the hourly pay required for an in-house paralegal to complete this work (and as a result, frees up paralegals to complete higher-value tasks.

### INITIAL USE CASES FOR AI AGENTS

| | |
|---|---|
| **Legal services** | EvenUp automatically builds demand packages for injury lawyers, allowing them to settle cases faster and more economically, and to sometimes win higher settlements than when all aspects of the workflow are done manually. |
| **Medical billing** | SmarterDx automates clinical documentation integrity (CDI) specialist work in hospital inpatient departments, analyzing 100% of the data contained in every patient chart to capture revenue that might otherwise be lost. |
| **Third-party insurance administration (TPAs)** | Reserv automates the insurance claim process and delivers better data and insights to carriers, MGAs, and other partners. Reserve is able to compete with and often replace legacy TPAs like Sedgwick and Crawford by delivering superior service through the use of generative AI in addition to talent. |

## Early pricing model examples of emergent Vertical AI leaders

Among many breakout Vertical AI businesses, we're seeing AI founders fully embrace usage and output-based pricing to align value capture with value creation. This approach of tying pricing to specific outcomes delivered means an easily quantifiable ROI for customers that can be benchmarked to their existing spend. In addition to this output-based pricing, vertical AI companies are often either putting this output into tiers or adding a base subscription fee to ensure a predictable baseline spend from each customer. This means that vertical AI companies can benefit from the predictability of subscription fees or tiered pricing while also capturing the upside of expanded use—a hybrid pricing represents an attractive model of value capture for vertical AI.

| Company | Company description | Pricing model |
|---|---|---|
| **DeepL** | AI-powered translation company that provides highly accurate and nuanced language translation services for businesses and individuals | Per user and editable file translation |
| **EvenUp** | AI-driven legal technology company that transforms the personal injury claims process by automating the creation of demand packages | Per demand package generated by AI |
| **Intercom** | AI customer communication platform that launched an AI agent that handles frontline customer support autonomously | FinAI agent with $0.99 per AI resolution; FinAI copilot with 10 free tickets per agent |
| **Zendesk** | A customer service and engagement platform that helps businesses build better relationships with their customers | Per automated ticket resolution |

## Up next: Our investment framework and founder advice

Vertical AI is a new frontier. While frameworks and strategies developed for horizontal AI and vertical software companies still have some utility for Vertical AI, we're also seeing a significant (and growing) number of unique opportunities, challenges, and trends that founders will need to understand and contend with in order to build innovative and valuable vertical AI products and services, successfully bring them to market, and maintain defensibly in the face of competition from new AI upstarts as well as AI features from vertical incumbents who aren't asleep at the wheel. In the fourth and final installment of our Vertical AI Roadmap, we'll share our initial investment framework and founder advice focused on four core pillars of vertical AI products and businesses: functional value, economic value, competitive dynamics, and defensibility.

PART IV

# Ten principles for building strong vertical AI businesses

We share our investment framework and essential advice for founders on how to drive early-stage value and defensibility.

A vertical AI business has many of the same building blocks as any great technology company. While essential lessons and frameworks that have helped founders create industry-defining vertical software businesses still apply, they're only part of the equation for those building the next vertical AI giant.

Vertical AI founders and teams are uncovering a bounty of largely untapped opportunities, and with these opportunities come novel considerations, challenges, and risks that warrant founders' attention.

Vertical AI is in a very early chapter, but the good news for founders is that blueprints of success do exist. There's a cohort of emergent vertical AI leaders whose founding predates the dawn of the AI era, such as EvenUp (founded in 2019), Subtle Medical (2017), Abridge (2018), and Fieldguide (2020) — companies led by prescient innovators who have identified areas where AI, with the right capabilities, could solve problems and drive ROI where past software companies couldn't.

Through our partnership and analysis of some of the most valuable, innovative, and defensible vertical AI solutions to date, we've formulated initial guiding principles for building strong vertical AI businesses from the ground up.

In the fourth and final installment of our Vertical AI Roadmap, we share ten key principles focused on strengthening the functional value, economic value, competitive position, and defensibility of vertical AI products, services, and businesses. Plus, we offer founders advice on applying these high-level principles to take practical action starting at the earliest stages of company-building.

## Our Vertical AI investment framework

| Bessemer Venture Partners | VALUE | | DEFENSIBILITY | |
|---|---|---|---|---|
| | Functional value | Economic value | Competitive dynamics | Defensibility |
| Good | Introduces new feature that demos well | Productivity boost | SaaS incumbent or weak AI upstarts | Execution speed as moat |
| Better | Introduces new feature with hard ROI, in particular, for a core business workflow | OpEx cost reduction | Adjacent competitor or sleepy incumbent | Complex product with early/weak data moat |
| Best | Tackles an end-to-end workflow with "LLM magic" not previously possible in absence of LLMs | Revenue boost or productivity improving for high value labor | No legitimate modern incumbent | Complex product with true moat around data and/or best-in-class multimodality |

# Guiding principles for building vertical AI businesses

## FUNCTIONAL VALUE

### 1. Determine which workflows your target customers want to automate.

Core and supporting workflows across different sectors will have varying propensities for automation. However, a workflow's propensity for automation — supporting or not — isn't the only important consideration when deciding what to build. Customers will inevitably have different degrees of interest in automation or have strict, specific requirements for automation, including in cases where builders view a given workflow as a perfect use case for automation.

Sometimes, these preferences or requirements can be addressed in the design of your product. For example, a dental office might want to set the procurement of medical supplies on autopilot if the order is below a certain cost but still want to have a human review larger purchases. An AI procurement solution can bake in that flexibility by completely automating certain orders while bringing a human in the loop for others. As another example, a law firm might be comfortable completely automating its client payment workflows, but, when it comes to a core workflow like writing legal briefs, they use human-in-the-loop-feedback to create the final output — such as creating an initial draft — because of potential pushback from the end client or a desire to maintain control over the final product.

As always, thorough market and user research is key. For example, in healthcare, Abridge and others have seen a swift adoption of AI solutions that address administrative workflows, and that's in large part because clinicians want to automate away administrative tasks such as scribing. And yet, while there's also interest in current applications of multimodal AI for diagnostics, some solutions have seen less adoption because healthcare's payment models have lagged behind its industry innovation.

The takeaway? Just because something can be automated doesn't mean it should be or that the conditions are right, given the possibility of industry-specific bottlenecks and other speedbumps. Your customer's guidance in where and how they want to use AI should always be your North Star.

### 2. Steer clear of easily commoditized applications.

Your vertical AI solution needs to differentiate itself in terms of what its core capabilities are — not just how well it executes those capabilities.

Certain solutions, such as data extraction and verification, will likely become table stakes, and building solutions with these kinds of easily replicable core value propositions will leave companies open to significant competitive pressure. For instance, we've seen increased accounts receivable and payable (AR/AP) automation solutions within financial services. While AI features for data matching and invoice reconciliation may offer some value, a solution that integrates these capabilities within a workflow tool that also creates tailored payment and membership plans for end customers would be more powerful and defensible — especially in the case of a vertical solution with sector-specific workflows.

For example, Sixfold, a generative AI leader for insurance underwriters, is designed to enhance underwriting capacity, accuracy, and transparency. Its solution is embedded as an API or plug-in within existing policy administration systems (PAS), eliminating the need for insurers to overhaul legacy systems or re-platform their workbench. This seamless integration allows underwriters to effortlessly introduce Sixfold's capabilities directly into their daily workflows.



### 3. Look for opportunities for AI to do work impossible for humans to do.

AI has been touted for its ability to take over rote and mundane tasks and workflows from employees so they have more time for the specialized work that they were trained and hired to do. However, the vertical AI products that provide the greatest value are also supplying teams with capabilities that humans don't have. One of the most common capabilities unique to AI is its ability to operate at scale.

For example, Rilla, an AI company in the home services vertical, provides coaching to sales representatives by transcribing and analyzing their in-person interactions with customers and then delivering tailored feedback and advice on improving performance. Without Rilla, a sales manager would have to physically accompany reps to onsite visits, and that manager's understanding of performance and best practices would always be limited by the number of time-consuming "ride-alongs" they could complete. Rilla, on the other hand, can audit high volumes of conversational data from reps across the company, meaning that the coaching it provides to sales reps is based on orders of magnitudes more data than a human could ever review individually.

There's a reason why certain industries like sales and marketing, services, and legal lend themselves to the magic of AI and LLMs: success in these categories requires both producing and deriving insights from large volumes of written text and, in some cases, audio recordings — time-consuming work that AI solutions can take the first pass, or even take over completely.

## ECONOMIC VALUE

### 4. The most valuable solutions unlock revenue.

Demonstrating your AI solution's hard ROI (either dollars saved or gained) for customers can expedite sales cycles considerably and increase customer retention. Solutions that improve profit margins by driving new revenue and/or reducing spend not only become an integral part of a customer's workflow but also of the performance of that customer's business.

One type of hard ROI is dollars saved via cost containment. For example, Abridge, which automates note-taking for doctor-patient interactions (among other things), has been able to show how much its platform is increasing doctor job satisfaction and, as a result, doctor retention. By driving retention, Abridge substantially mitigates the cost of recruiting and training physicians — costs which often amount to millions, if not tens of millions, annually.

On top of cost containment, Abridge contributes to new revenue by freeing up one to two hours daily per physician. These extra hours enable physicians to see more patients, directly boosting the hospital's operational efficiency and generating revenue that wouldn't be generated in the absence of the solution. Abridge's detailed transcripts and summaries of each patient visit also prevent revenue leakage by ensuring comprehensive coding and billing.



EvenUp is another vertical AI company that directly increases its customers' revenue by freeing up time, in this case, for legal professionals. Using a combination of the latest LLM technology and proprietary models (PiAi), EvenUp generates demand packages for personal injury law firms in a fraction of the time it would take a paralegal (who needs days to gather data from the client, collate hundreds of documents, extract data from medical and police reports, etc.). Because EvenUp's legal operations team reviews each letter, law firms can maintain a high-quality standard and still drastically reduce (or eliminate) the time that their team spends on-demand packages. This additional time allows firms to take on more cases which increases revenue.

### 5. Novel business models can unlock new opportunities.

Vertical AI solutions are being delivered and priced in novel ways, which has opened up new verticals for automation — verticals where there wasn't sufficient TAM to build a traditional software business. This is especially true of the AI-enabled services model, where AI automates tasks to deliver potentially faster, cheaper, and more consistent services than legacy service providers.

Because of the lower cost structure inherent to automation, in 2024 we see, on average, a ~56% gross margins on a 1.6x average burn ratio in our vertical AI portfolio, which includes a significant number of services companies. Historically, services businesses have been hard to make profitable because of the costs of specialized workers, but with AI, specialized workers can play a more limited role, focused on quality control and reinforcement learning from human feedback (RLHF).

Some AI services products will be better delivered with human QA support, while others will be equally or more effective with an internal AI-powered product as the core service offering. The best way to deliver your AI solution will ultimately depend on the customer you're serving and their comfort level engaging products and services offering various degrees of automation, as well as other considerations.

## COMPETITIVE DYNAMICS

### 6. Build for overlooked categories and workflows.

Compared to broader horizontal categories like sales or marketing where scaled and well-resourced competitors (such as Salesforce or ADP) already exist, founders will typically find less incumbent pressure in vertical categories — especially under-the-radar verticals where few innovative companies are operating. While it's ideal to claim the first-mover advantage in a greenfield environment, most vertical categories will already have at least one incumbent. However, when incumbents are stretched thin or slow to integrate AI, fast-moving startups can gain a competitive edge by building superior, high-ROI AI products and services that address workflows where automation is a valuable but non-obvious solution, either because it's an unconventional use case or one where it's difficult to execute with AI successfully.

### 7. Serve customers with nuanced needs.

Vertical AI companies can further differentiate by targeting customers within overlooked categories that have particularly complex requirements and nuanced needs that can't be easily met by an AI solution. For example, an AI startup serving banks or government contractors would need to build sector-specific security and compliance tooling to sell to these buyers, which would add another vector of defensibility to the solution. To mitigate LLM commoditization risk, we will likely begin to see foundational model players, such as OpenAI and Anthropic, go off to build models that support very specific and specialized use cases for customers in industries like these.

### 8. Models aren't a reliable moat — but multimodality can be.

As model infrastructure costs continue to plummet, models will cease to be a moat, and early vertical AI founders will need to ask themselves: "Why is the product we're building with AI going to be superior to what can be built with publicly available models and data?"

Constructing your technical architecture to address the specific problem you're solving is a start; for example, fine-tuning LLMs to better reflect a customer's writing style or using Retrieval-augmented generation (RAG) to better execute information retrieval. (We believe using RAG for industry-specific datasets is a foundational level of defensibility.)

But we expect that additional layers of defensibility will be found in solutions that have the capabilities to address more complex — and particularly multimodal — workflows. For example, Jasper, the AI marketing platform, also went multimodal when they acquired Clickdrop for image generation, and then built additional enterprise marketing capabilities, including Brand Voice, and Campaign Calendaring.

## DEFENSIBILITY

### 9. Focus on modularity and scalability in your model stack.

Unlike traditional SaaS businesses, where companies built on similar permutations of a standard technology stack, vertical AI companies need to build with bespoke, best-in-class infrastructure and LLMs in lockstep with the pace of innovation in AI — or else risk getting left behind. Teams can stay nimble by building internal capabilities to flexibly combine open-source, fine-tuned, and proprietary models so that your tech achieves the best outcome for your customer.

This approach allows you to move and execute quickly in a rapidly developing infrastructure landscape and take a plug-and-play approach to your technical architecture. If you can get the same quality result by taking an open-source LLM and fine-tuning it, it's not worth building a model from scratch (especially if you would otherwise end up with technical debt when the market or tech changes). This approach will also allow you to put your resources towards what matters most: getting a superior product to your customer. It's very unlikely that your customers will care about the details of the underlying model as long as the solution that you provide gets the job done better than any alternative available to them.



Again, a great example of a product built for flexibility is Jasper, the platform that sits in the nucleus of the marketing tech stack and acts as an "AI brain" to help users formulate, design, and execute initiatives across all marketing specialties. The Jasper team has architected a modular platform that uses multiple LLMs. It runs marketing inputs through several LLMs depending on the customer need, model performance, and cost. For example, if Claude 3.5 outperforms GPT-4 for a particular use case, Jasper's product architecture is flexible enough to support interchangeable model infrastructure.

### 10. Don't over-index on the quantity of data. It's the quality that counts.

Much has been said about the importance of proprietary data sets as a key to defensibility. As an early-stage startup in a new category, builders likely can't get the volume they're looking for. However, high-quality data (of any amount) will have a compounding effect, benefitting the company more and more over time.

For example, in the early days of EvenUp, the team dramatically and consciously invested in legal operations to have humans-in-the-loop to review all demand letters; this was a case where data scale or mass didn't matter as much as data quality and honing the model with feedback over time to improve the product.

What matters most at the early stage is building a solid, high-ROI initial product that's meeting your core customer's pain points and flying off the shelves, so getting initial quality data for models needs to happen without compromising fast-paced execution, especially if you're trying to be first to market. With usage at scale, proprietary data will come, and then that data can be used to strengthen your product.

As another example, an AI-powered marketplace that automates the creation of requests for proposals (RFPs) for vendors could use historical and market data that comes as buyers and sellers get added to the platform, and that data could be used to produce RFPs that are more likely to win business. But when your marketplace is just getting started, your ability to surface good business opportunities for vendors and surface quality vendors for buyers will matter most.

## Defensibility

What continues to astonish us at Bessemer is the sheer scale of vertical AI businesses and their potential to surpass their vertical SaaS predecessors. This massive growth is driven, in part, by Vertical AI's ability to access labor and services budgets rather than being confined to traditional IT software budgets.

Vertical AI solutions are reimagining systems of record as systems of work, offsetting future headcount growth, and amplifying productivity by equipping employees with entirely new capabilities. Unlike vertical SaaS, which typically captures only a fraction of a Fortune 500's IT spend, Vertical AI taps into the labor line of a P&L, accessing exponentially larger budgets.

Emergent leaders in this category have already demonstrated how Vertical AI can transform underserved, service-heavy industries by automating essential workflows and delivering value where prior software solutions fell short. These companies are tangible case studies for vertical AI businesses yet to be started, and through them, aspiring founders can learn powerful lessons that they likely couldn't find in a book doled out in business school.



Our entire Vertical AI roadmap stems from deep partnerships and hundreds of conversations with entrepreneurs, as well as extensive research and decades of experience investing in vertical software. However, we must not mistake the existence of today's frameworks and strategies as a sign that the category is mature. Vertical AI is still in its infancy. Opportunities remain vast, and the landscape will only grow more dynamic as advancements in generative AI and foundation models open new frontiers.

Many of tomorrow's Vertical AI titans are being built today — but countless others are still ideas waiting to be imagined and brought to life.

**Venture insights that matter**

bvp.com/subscribe