

The System That Will Win Enterprise AI

As we start Q2, I want to begin where we always should: with our customers. I have been spending time with leaders across our largest enterprises, most influential startups, and key venture firms. The message is clear. People are excited about what we are building, and they want a deeper view into our roadmap so they can plan with confidence and stay ahead of the market.

Enterprise AI is entering a more mature phase. Raw capability still matters, but it is no longer enough. Customers want fit: how well AI plugs into their workflows, knowledge, controls, and day-to-day operations, and how effectively it can be deployed, trusted, and improved over time. They want a system they can trust and build on.

We are building that system: the best models for work, a platform for agents, deep integration with business context, and the ability to deploy and improve at scale. And customers are validating that direction in the clearest possible way. Multi-year, multi-product, nine-figure deals are rising, and existing customers are expanding as they standardize on our capabilities across more of their organizations.

I am incredibly proud of how this team is showing up. We are earning trust through the depth, quality, and care we bring to the work. The opportunity ahead is massive, and our biggest constraint right now is not demand. It is capacity. That is why talent remains a top priority in Q2. We will keep hiring deliberately, keep the bar high, and keep building a team that matches the

excellence our customers expect from us and we expect from each other.

We have everything we need to extend our lead from here. We have the compute. We have the products. We have the customer pull. This is the moment to lean in and make the case, clearly and confidently, that OpenAI is the platform enterprises should trust to build, deploy, and scale with.

Here are five customer-backed priorities I want us to focus on.

1. Win the model layer for work

Enterprises buy business outcomes. They pay for models that help employees write faster, analyze better, code more productively, support customers more effectively, and make higher-quality decisions. They pay for higher revenue per employee, faster cycle times, lower support costs, and better execution.

Spud is an important step in the intelligence foundation for the next generation of work. Early feedback from our customers is very positive. Spud is not only our smartest model yet, but it also delivers on everything that matters for high-value professional work: stronger reasoning, better understanding of intent and dependencies, better follow-through and more reliable output in production.

Better model performance lifts the rest of the stack. Spud will make all of our key products significantly better. It expands the workflows we can own and gives customers another reason to consolidate around us. This

is our iterative deployment strategy in practice: push the frontier, deploy it into real products, learn from real usage, and compound those lessons into better systems on the path to the super app.

Our compute advantage sets us up to deliver continuous leaps in capability. Customers already feel it in real product terms: higher token limits, lower latency, and more reliable execution of complex workflows. Every step forward in compute lets us train stronger models, serve more demand, and lower the cost per unit of intelligence. That is durable business leverage.

2. Win the agent platform layer

The market has moved from prompts to agents. That shift is a massive opportunity for us.

Customers want systems that can reason, use tools, operate across workflows, and perform reliably inside real business environments. That means orchestration, control, observability, security, integration, and governance.

Frontier allows us to own the platform layer. We need to position Frontier as the default platform for enterprise agents – the core intelligence layer enterprises use to build, deploy, manage, and scale systems.

This is where our advantage can compound. Frontier ties model intelligence directly to agent performance. As our models improve, the platform gets more valuable. As the platform gets embedded, switching costs rise. As customers run more workflows through the system,

OpenAI becomes harder to replace and more central to how work gets done.

That is how we move from product vendor to operating infrastructure.

3. Expand the market through Amazon

Our Microsoft partnership has been foundational to our success. But it has also limited our ability to meet enterprises where they are – for many that’s Bedrock.

Since we announced the partnership at the end of February, inbound demand from our customers for this offering has been frankly staggering. We are firing on all cylinders to establish this as a scaled distribution channel.

The Amazon Stateful Runtime Environment matters because it expands access and upgrades the product surface at the same time. By enabling memory, context, and continuity across interactions, we move beyond stateless model access toward systems that can operate reliably over time and across complex business processes.

This will expand our market in three ways: 1. It lowers adoption friction for AWS-native customers. 2: It strengthens our position with regulated and security-sensitive buyers by running inside their AWS environment and existing governance model. 3. It further integrates our platform from model access to production runtime for long-running, multi-step agents.

4. Sell the full AI-native stack

Customers want a platform not point solutions. That's what we have: ChatGPT for Work is the front door for knowledge work. Codex is the system for software and agentic development. The API is the engine for embedded intelligence inside customer products and workflows. Frontier is the agent platform. The Amazon runtime extends our reach into production-grade, stateful execution.

That breadth is a major strategic advantage because customers do not all start in the same place. Some start with employees. Some start with developers. Some start with internal systems. Some start with external products. Our job is to meet them wherever they enter and then expand them across the full stack.

This is the flywheel we should be building around: better models drive more usage, more usage drives deeper integration, deeper integration drives multi-product adoption, and multi-product adoption makes us harder to replace.

We should stop thinking like a company with separate product lines. We should think like a platform company with multiple entry points and one integrated enterprise offering.

5. Own deployment

The biggest bottleneck in enterprise AI is no longer whether the technology works. It is whether

companies can deploy it successfully and at scale.

DeployCo gives us the chance to turn product demand into repeatable enterprise transformation. It will be a deployment engine that helps companies prove value faster, reduce risk, and scale adoption across the organization.

This can become a force multiplier across everything else we are building. It helps customers move faster. It sharpens our feedback loops. It surfaces repeatable deployment patterns. It improves product, sales, and customer success all at once. And, alongside our Frontier Alliance partners, it gives us a serious path to scale execution across the market.

The companies that win enterprise AI will not just have the best models. They will have the best ability to get those models deployed into real workflows, inside real organizations, with real measurable value. We should be the best in the world at that.

A note on the competitive landscape

The market is as competitive as I have ever seen it. I believe that is ultimately a good thing. It means the opportunity is immense and important. However, there is no question it can be noisy, volatile and distracting at times. Competition inspires us and will make us all better and most importantly our customers will feel that benefit. To that point, as you have not heard me say many times, the number one focus should be spending time with our customers. When we spend time with our customers, listening to what their problems and

ambitions are, focusing on how we can invest in them and help, everything else gets quiet and comes into focus.

With that all being said, here are a few things worth keeping in mind, especially on Anthropic.

- Their story is built on fear, restriction, and the idea that a small group of elites should control AI. Our positive message will win over time: build powerful systems, put in the right safeguards, expand access, and help people do more.
- Their strategic misstep to not acquire enough compute is showing up in the product. Customers feel it through throttling, weaker availability, and a less reliable experience. We saw the exponential compute curve earlier, acted on it faster, and now have a real structural advantage.
- Their coding focus gave them an early wedge. But you do not want to be a single-product company in a platform war. As AI spreads beyond developers into every team, workflow, and industry, that narrowness can become a real liability.
- Their stated run rate is inflated. They use accounting treatment that makes revenue look bigger than it is, including grossing up rev share with Amazon and Google. Our analysis shows that this overstates their run rate by roughly \$8 billion (at the current \$30 stated). We report Microsoft revshare net, which is more inline with standards we would be held to as a public company.

Let's Go Build

Finally, one of the best things about the work we do is the people we get to do it with. I am so proud of this company and our team. It is a privilege to work with all of you and to be alive at this moment in the epicenter of the future. Lets all stay focused, work as one team and operate at the highest level of excellence and row in the same direction.

The market is ours to win, let's execute accordingly.