OpenAI

# AI in the Enterprise

Lessons from seven frontier companies

# Contents

# A new way to work

As an AI research and deployment company, OpenAI prioritizes partnering with global companies because our models will increasingly do their best work with sophisticated, complex, interconnected workflows and systems.

We're seeing AI deliver significant, measurable improvements on three fronts:

| | | |
|---|---|---|
| 01 | **Workforce performance** | Helping people deliver higher-quality outputs in shorter time frames. |
| 02 | **Automating routine operations** | Freeing people from repetitive tasks so they can focus on adding value. |
| 03 | **Powering products** | By delivering more relevant and responsive customer experiences. |

But leveraging AI isn't the same as building software or deploying cloud apps. The most successful companies are often those who treat it as a new paradigm. This leads to an experimental mindset and an iterative approach that gets to value faster and with greater buy-in from users and stakeholders.



# Our approach: iterative development

OpenAI is organized around three teams. Our **Research Team** advances the foundations of AI, developing new models and capabilities. Our **Applied Team** turns those models into products, like ChatGPT Enterprise and our API. And our **Deployment Team** takes these products into companies to address their most pressing use cases.

We use iterative deployment to learn quickly from customer use cases and use that information to accelerate product improvements. That means shipping updates regularly, getting feedback, and improving performance and safety at every step.

The result: users access new advancements in AI early and often—and your feedback shapes future products and models.

# Seven lessons for enterprise AI adoption

| 01 | **Start with evals** | Use a systematic evaluation process to measure how models perform against your use cases. |
|----|---------------------|------------------------------------------------------------------------------------------|
| 02 | **Embed AI in your products** | Create new customer experiences and more relevant interactions. |
| 03 | **Start now and invest early** | The sooner you get going, the more the value compounds. |
| 04 | **Customize and tune your models** | Tuning AI to the specifics of your use cases can dramatically increase value. |
| 05 | **Get AI in the hands of experts** | The people closest to a process are best-placed to improve it with AI. |
| 06 | **Unblock your developers** | Automating the software development lifecycle can multiply AI dividends. |
| 07 | **Set bold automation goals** | Most processes involve a lot of rote work, ripe for automation. Aim high. |

Let's drill down into each of these, with customer stories as examples.

# Start with evals

**How Morgan Stanley iterated to ensure quality and safety**

As a global leader in financial services, Morgan Stanley is a relationship business. Not surprisingly, there were some questions across the business about how AI could add value to the highly personal and sensitive nature of the work.

The answer was to conduct intensive evals for every proposed application. An eval is simply a rigorous, structured process for measuring how AI models actually perform against benchmarks in a given use case. It's also a way to continuously improve the AI-enabled processes, with expert feedback at every step.

**How it started**

Morgan Stanley's first eval focused on making their financial advisors more efficient and effective. The premise was simple: If advisors could access information faster and reduce the time spent on repetitive tasks, they could offer more and better insights to clients.

**They started with three model evals:**

| | | |
|---|---|---|
| 01 | **Language translation** | Measuring the accuracy and quality of translations produced by a model. |
| 02 | **Summarization** | Evaluating how a model condenses information, using agreed-upon-metrics for accuracy, relevance, and coherence. |
| 03 | **Human trainers** | Comparing AI results to responses from expert advisors, grading for accuracy and relevance. |

These evals—and others—gave Morgan Stanley the confidence to start rolling the use cases into production.

**How it's going**

Today, 98% of Morgan Stanley advisors use OpenAI every day; access to documents has jumped from 20% to 80%, with dramatically reduced search time; and advisors spend more time on client relationships, thanks to task automation and faster insights.

*The feedback from advisors has been overwhelmingly positive. They're more engaged with clients, and follow-ups that used to take days now happen within hours.*

**Kaitlin Elliott**
Head of Firmwide Generative AI Solutions



To find out more, watch Morgan Stanley: Shaping the Future of Financial Services and ask us about our Eval Frameworks.

# Evals defined

Evaluation is the process of validating and testing the outputs that your models produce. Rigorous evals lead to more stable, reliable applications that are resilient to change. Evals are built around tasks that measure the quality of the output of a model against a benchmark—is it more accurate? More compliant? Safer? Your key metrics will depend on what matters most for each use case.

# Embed AI into your products

**How Indeed humanizes job matching**

When AI is used to automate and accelerate tedious, repetitive work, employees can focus on the things only people can do. And because AI can process huge amounts of data from many sources, it can create customer experiences that feel more human because they're more relevant and personalized.

Indeed, the world's No. 1 job site, uses GPT-4o mini to match job seekers to jobs in new ways.

**The power of why**

Making great job recommendations to job seekers is only the start of the Indeed experience. They also need to explain to the candidate *why* this specific job was recommended to them.

Indeed uses the data analysis and natural language capabilities of GPT-4o mini to shape these 'why' statements in their emails and messages to jobseekers. Using AI, the popular 'Invite to Apply' feature also explains why a candidate's background or previous work experience makes the job a good fit.

The Indeed team tested the previous job matching engine against the GPT-powered version with the new, customized context. The performance uplift was significant:

**A 20% increase** in job applications started

**A 13% uplift in downstream success**—not only were more candidates likely to apply, but employers were more likely to hire them.

With Indeed sending over 20 million messages a month to job seekers—and 350 million visitors coming to the site every month—these increases scale up to significant business impact.

But scaling up also meant using more tokens. To increase efficiency, OpenAI and Indeed worked together to fine-tune a smaller GPT model that was able to deliver similar results with 60% fewer tokens.

Helping job seekers find the right jobs—and understanding *why* a given opportunity is right for them—is a profoundly human outcome. Indeed's team has used AI to help connect more people to jobs, faster—a win for everyone.

> *We see a lot of opportunity to continue to invest in this new infrastructure in ways that will help us grow revenue.*
>
> **Chris Hyams**
> CEO

# Start now and invest early

**How Klarna benefits from AI knowledge compounding**

AI is rarely a plug-and-play solution—use cases grow in sophistication and impact through iteration. The earlier you start, the more your organization benefits from compounding improvements.

Klarna, a global payments network and shopping platform, introduced a new AI assistant to streamline customer service. Within a few months, the assistant was handling two-thirds of all service chats—doing the work of hundreds of agents and cutting average resolution times from 11 minutes to just 2. The initiative is projected to deliver $40 million in profit improvement, all while maintaining satisfaction scores on par with human support. These results didn't happen overnight. Klarna achieved this performance by continuously testing and refining the assistant.

Just as importantly, 90% of Klarna's employees now use AI in their daily work. Growing organization-wide familiarity with AI has enabled Klarna to move faster, launch internal initiatives more efficiently, and continuously refine the customer experience. By investing early and encouraging broad adoption, Klarna is seeing AI's benefits compound—driving returns across its business.

AI in the Enterprise

*This AI breakthrough in customer interaction means superior experiences for our customers at better prices, more interesting challenges for our employees, and better returns for our investors.*

**Sebastian Siemiatkowski**
Co-Founder and CEO
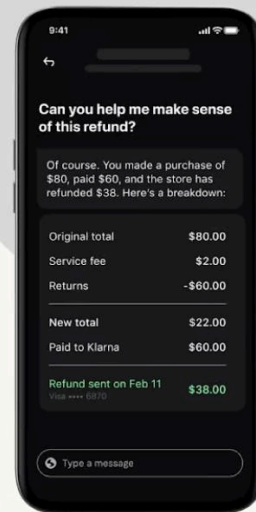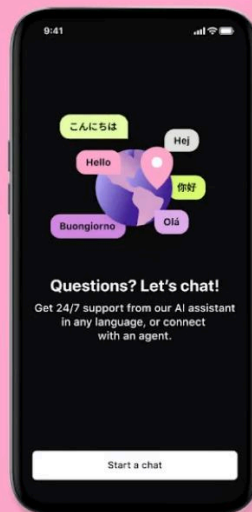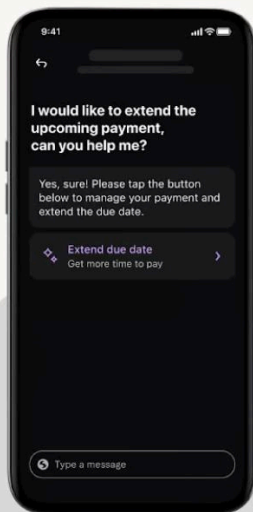
# Customize and fine-tune your models

**How Lowe's improves product search**

Enterprises seeing the most success from AI adoption are often the ones that invest time and resources in customizing and training their own AI models. OpenAI has invested heavily in our API to make it easier to customize and fine-tune models—whether as a self-service approach or using our tools and support.

We worked closely with Lowe's, the Fortune 50 home improvement company, to improve the accuracy and relevance of their ecommerce search function. With thousands of suppliers, Lowe's often has to work with incomplete or inconsistent product data.

The key is in accurate product descriptions and tagging. But it also requires an understanding of how shoppers search, a dynamic that changes across product categories. That's where fine-tuning comes in.

By fine-tuning OpenAI models, the Lowe's team was able to improve product tagging accuracy by 20%—with error detection improving by 60%.

*Excitement in the team was palpable when we saw results from fine-tuning GPT 3.5 on our product data. We knew we had a winner on our hands!*

**Nishant Gupta**
Senior Director, Data, Analytics and Computational Intelligence



Product Note: OpenAI has launched Vision Fine-Tuning to further improve ecommerce search and address challenges in medical imaging and autonomous driving.

**What is fine-tuning?**

# If a GPT model is a store-bought suit, fine-tuning is the tailored option—the way you customize the model to your organization's specific data and needs.

**Why it matters:**

**Improved accuracy**

By training on your unique data—such as product catalogs or internal FAQs—the model delivers more relevant, on-brand results.

**Domain expertise**

Fine-tuned models better understand your industry's terminology, style, and context.

**Consistent tone and style**

For a retailer, that could mean every product description stays true to brand voice; for a law firm, it means properly formatted citations, every time.

**Faster outcomes**

Less manual editing or re-checking means your teams can focus on high-value tasks.

# Get AI in the hands of experts

**BBVA takes an expert-led approach to AI**

Your employees are closest to your processes and problems and are often the best-placed to find AI-driven solutions. Getting AI into the hands of these experts can be far more powerful than trying to build generic or horizontal solutions.

BBVA, the global banking leader, has more than 125,000 employees, each with a unique set of challenges and opportunities. They decided to get AI into the hands of employees—working closely with Legal, Compliance, and IT Security teams to ensure responsible use. They rolled out ChatGPT Enterprise globally, then let people discover their own use cases.

"Normally, in a business like ours, building even a prototype requires technical resources and time," says Elena Alfaro, Head of Global AI Adoption at BBVA. "With custom GPTs, anyone can create apps to solve unique problems—it's very easy to start."

In five months, BBVA employees created over 2,900 custom GPTs—some of which reduce project and process timelines from weeks to hours. The impact was felt across many disciplines and departments:

| | |
|---|---|
| **The Credit Risk team** | Uses ChatGPT to determine creditworthiness faster and more accurately. |
| **The Legal team** | Uses it to answer 40,000 questions a year on policies, compliance, and more. |
| **The Customer Service team** | Automates the sentiment analysis of NPS surveys. |

And the wins continue to spread across Marketing, Risk Management, Operations, and beyond. All because they got AI in the hands of the people who know how to apply it in their own disciplines.

*We consider our investment in ChatGPT an investment in our people. AI amplifies our potential and helps us be more efficient and creative.*

**Elena Alfaro**
Head of Global AI Adoption



Product Note: With deep research, ChatGPT can do work independently. Give it a prompt, and it can synthesize hundreds of online sources to create comprehensive, PhD-level reports. This unlocks employee productivity and gives them access to deep, detailed research on any topic in minutes. In an internal evaluation by experts across domains, deep research saved an average of 4 hours per complex task.

For more detail, watch <u>BBVA puts AI into the hands of every team</u>.

# Unblock your developers

**Mercado Libre builds AI programs faster and more consistently**

Developer resources are the main bottleneck and growth inhibitor in many organizations. When engineering teams are overwhelmed, it slows innovation and creates an insurmountable backlog of apps and ideas.

Mercado Libre, Latin America's largest ecommerce and fintech company, partnered with OpenAI to build a development platform layer to solve that. It's called Verdi, and it's powered by GPT-4o and GPT-4o mini. Today, it helps their 17,000 developers unify and accelerate their AI application builds.

Verdi integrates language models, Python nodes, and APIs to create a scalable, consistent platform that uses natural language as a central interface. Developers now build consistently high-quality apps, faster, without having to get into the source code. Security, guardrails, and routing logic are all built in.

As a result, AI app development has accelerated dramatically, helping Mercado Libre employees do amazing things, including:

| | |
|---|---|
| **Improving inventory capacity** | GPT-4o mini Vision tags and completes product listings, allowing Mercado to catalog 100x more products. |
| **Detecting fraud** | Evaluating data on millions of product listings each day, improving fraud detection accuracy to nearly 99% for flagged items. |
| **Customizing product descriptions** | Translating product titles and descriptions to adapt to nuanced Spanish and Portuguese dialects. |
| **Increasing orders** | Automating review summaries to help users quickly grasp product feedback. |
| **Personalizing notifications** | Tailoring push notifications to drive higher engagement and improve product recommendations. |

Next up: using Verdi to improve logistics, reduce late deliveries, and take on more high-impact tasks across the organization.

*We designed our ideal AI platform using GPT-4o mini, with a focus on lowering cognitive load and enabling the entire organization to iterate, develop, and deploy new, innovative solutions.*

**Sebastian Barrios**
SVP of Technology
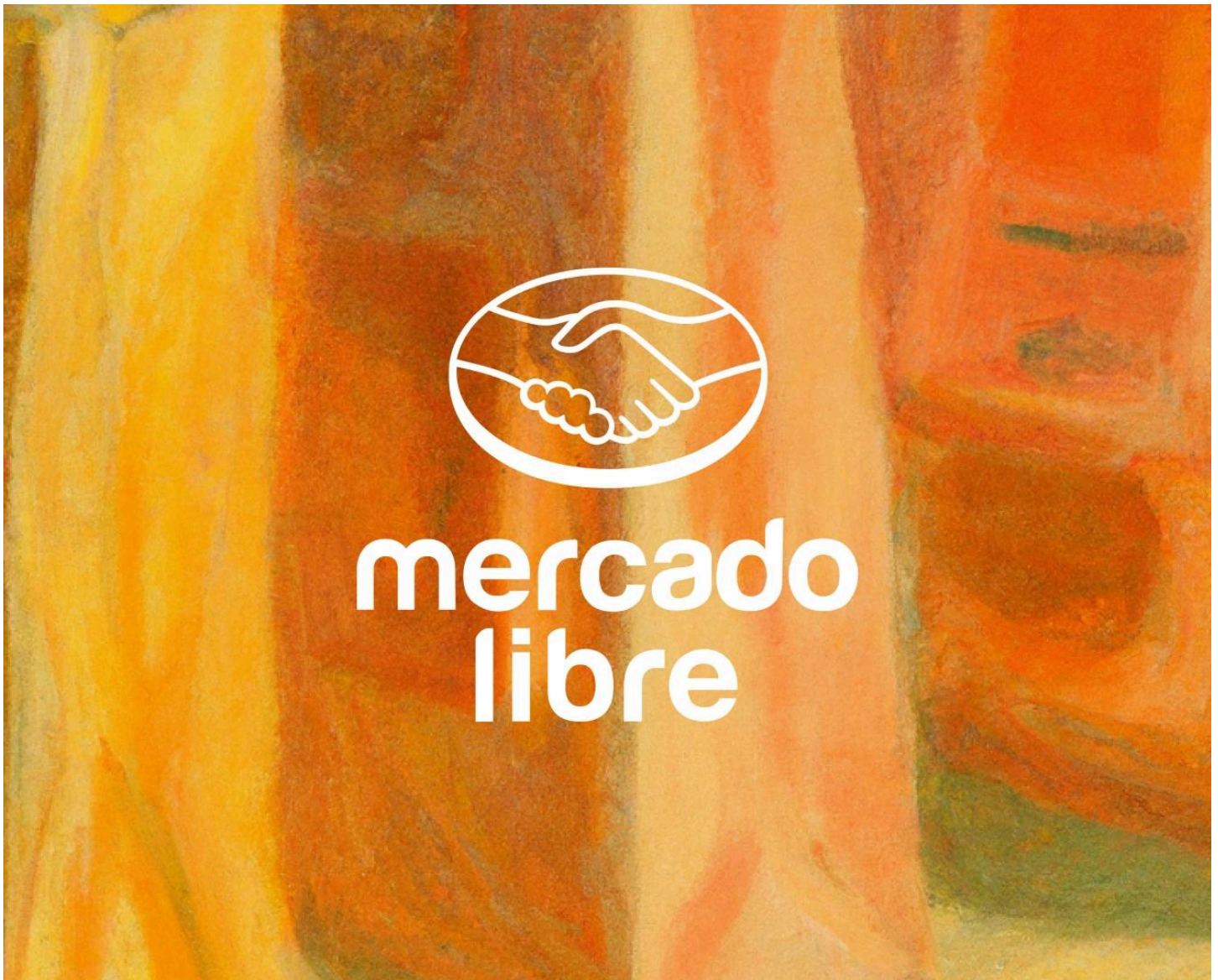
# Set bold automation goals

**How we automate our own work at OpenAI**

At OpenAI, we live with AI every day, so we're often spotting new ways to automate our own work.

An example: Our support teams were getting bogged down, spending time accessing systems, trying to understand context, craft responses, and take the right actions for customers.

So we built an internal automation platform. It works on top of our existing workflows and systems to automate rote work and accelerate insight and action.

Our first use case: working on top of Gmail to craft customer responses and trigger actions. Using our automation platform, our teams can instantly access customer data and relevant knowledge articles, then incorporate the results into response emails or specific actions—such as updating accounts or opening support tickets.

By embedding AI into existing workflows, our teams are more efficient, responsive, and customer-focused. This platform handles hundreds of thousands of tasks every month, freeing people to do more high-impact work. Not surprisingly, the system is now spreading across other departments.

It happened because we set bold automation goals from the start, instead of accepting inefficient processes as a cost of doing business.

# Learning from each other

As the previous examples show, every business is full of opportunities to harness the power of AI for improved outcomes. The use cases may vary by company and industry but the lessons apply across all markets.

The common theme: AI deployment benefits from an open, experimental mindset, backed by rigorous evaluations, and safety guardrails. The companies seeing success aren't rushing to inject AI models into every workflow. They're aligning around high-return, low-effort use cases, learning as they iterate, then taking that learning into new areas.

The results are clear and measurable: faster, more accurate processes; more personalized customer experiences; and more rewarding work, as employees focus on the things people do best.

We're now seeing companies integrating AI workflows to automate increasingly sophisticated processes—often using tools, resources, and other agents to get things done.

We'll continue to report back from the front lines of AI to help guide your own thinking.

**Product Note: Operator**

Operator is an example of OpenAI's agentic approach. Leveraging its own virtual browser, Operator can navigate the web, click on buttons, fill in forms, and gather data just like a human would.

It can also run processes across a wide range of tools and systems—no need for custom integrations or APIs. Enterprises use it to automate workflows that previously required human intervention, such as:

> Automating software testing and QA using Operator to interact with web apps like a real user, flagging any UI issues.

> Updating systems of record on behalf of users, without technical instructions or API connections.

The result: end-to-end automation, freeing teams from repetitive tasks and boosting efficiency across the enterprise.

# The trusted AI enterprise platform

**Security and privacy at a glance**

For our enterprise customers, nothing is more important than security, privacy and control. Here's how we ensure it:

**Your data stays yours**

We don't use your content to train our models; your enterprise retains full ownership.

---

**Enterprise-grade compliance**

Data is encrypted in transit and at rest, aligned with top standards like SOC 2 Type 2 and CSA STAR Level 1.

---

**Granular access controls**

You choose who can see and manage data, ensuring internal governance and compliance.

---

**Flexible retention**

Adjust settings for logging and storage to match your organization's policies.

For more on OpenAI and security, visit our Security page or the OpenAI Security Portal.

# More resources

OpenAI for Business

OpenAI Stories

ChatGPT Enterprise

OpenAI and Safety

API Platform


OpenAI is an AI research and deployment company. Our mission is to ensure that artificial general intelligence benefits all of humanity.

# OpenAI