OpenAI

How enterprises are scaling AI from curiosity to compounding impact

Practical insights from European enterprise leaders in the field

PHILIPS BBVA SCANIA JETBRAINS Scout24 MIRAKL

Frontiers of AI

May 2026

Contents

Foreword

We all know AI model capabilities have accelerated fast. But many enterprises still struggle to translate that intelligence into reliable, production-grade outcomes.

That delta is what we describe as the capability gap: what AI can do is moving ahead of what most organizations can actually deploy. It shows up as pilots that don't scale, point solutions layered onto already fragmented systems, and teams that can't quite connect experimentation to operational impact. This is a leadership challenge, not a technical one.

At OpenAI our goal is simple: help organizations turn AI into measurable outcomes, guided by our north star of building the most capable AI technology that benefits everyone. We spoke with leaders from five companies on the forefront of AI adoption and deployment – Philips, BBVA, Mirakl, Scout24, Jetbrains and Scania – about what they're learning as they scale AI. This is a guide by and for leaders, built from those conversations, to help you move from experiments to durable impact.

The organizations that win with AI won't be the ones that tried it first. They'll be the ones that operationalized it best – turning intelligence into infrastructure and adoption into a durable advantage.

Sanj Bhayro
Managing Director , EMEA, OpenAI

Executive summary

In our first *Frontiers of AI* interview series, enterprise executives on the front lines of AI adoption shared a consistent insight: lasting impact comes not from model access, but from the early decisions leaders make about trust, ownership, and quality.

The organizations featured in this series did not treat AI as a rollout or a tooling upgrade. They treated it as an operating layer and leadership discipline. Trust was established early. AI was shaped around real workflows rather than novelty. Quality was defined, measured, and protected before scale.

These choices unlocked durable speed — and, over time, compounding impact.

This guide synthesizes what leaders across healthcare, finance, commerce, media, and software learned while moving AI from experimentation into production. It is not a set of predictions or promises. It is an evidence-based view of what has held true in practice.

A shift driven by trust, workflow design, and evidence at scale

Frontiers of AI explores how organizations at the leading edge of adoption are turning ambition into something operational.

Across Philips, BBVA, Mirakl, Scout24, JetBrains, and Scania, the industries differ. The challenges do not. In each case, leaders confronted the same questions: how to earn trust, how to redesign work, how to govern responsibly, and how to scale without losing quality.

The playbook is converging — not around technology choices, but around leadership decisions.

Five patterns that consistently enabled scale

01 Culture before tooling

At Philips, the question was never whether AI would matter. It was how to embed AI into everyday work across a 70,000-strong workforce in healthcare and technology.

Rather than positioning AI as a specialist capability, leaders focused on literacy and confidence. Senior leaders were trained first. Bottom-up ideas surfaced practical use cases. Teams were supported in moving deliberately from experimentation to real workflows.

By treating AI as an organizational capability — not a toolkit — Philips drove broad adoption aligned with patient care and clinical quality.

> "Turning AI from a niche skill into a company-wide capability takes more than technology; it takes culture, curiosity, and courage to experiment."
>
> **Patrick Mans**
> Head of Data Science & AI, Philips

02 Governance as an enabler

BBVA's experience shows what happens when governance is designed to enable speed rather than constrain it.

Security, legal, compliance, and IT were involved from the outset — not as reviewers at the end, but as partners in design. This foundation made it possible to move quickly without losing control, a non-negotiable requirement in a global, regulated bank.

Importantly, governance worked because it reinforced trust. Employees had confidence that AI could be used for real thinking and everyday problem-solving — not just "safe" or superficial tasks.

Early deployments focused on concrete daily workflows. In Peru, an internally built assistant adopted by more than 3,000 employees reduced average query handling time from roughly 7.5 minutes to around one minute. When AI meaningfully improved daily work, adoption followed.

What began as targeted experimentation has quickly expanded into a bank-wide capability, with ChatGPT Enterprise now rolling out to more than 120,000 employees globally.

> "Once you start using it, it's very sticky. You feel it helps you a lot."
>
> **Antonio Bravo**
> Global Head of Data, BBVA

03 Ownership over consumption

Mirakl moved its ambition beyond universal AI usage toward something more demanding: enabling teams to build.

Employees were given the tools and autonomy to create their own agents and redesign workflows with AI in mind. Adoption shifted from passive consumption to active ownership.

In practice, scale was driven less by central committees and more by embedded champions — people close to the work, empowered to turn experimentation into habit and local success into shared capability.

The impact was measurable:

70%

Faster creation

of internal technical documentation

37%

Customer support efficiency

while maintaining **96%** customer satisfaction

91%

Catalog onboarding acceleration

with roughly half the previous error rate

The lesson was not to deploy AI everywhere, but to use it deliberately as a system-level lever — while keeping human judgment firmly where it matters.

> "When you embrace AI, think about objectives — scale, timing, reach, impact — that are far greater than what you're used to planning for."
>
> **Adrien Nussenbaum**
> Co-Founder & Co-CEO, Mirakl

At Scania, this shift showed up in how AI was embedded directly into operational workflows — from engineering to service — rather than treated as a separate capability. Teams focused on applying AI where it could remove friction from complex, real-world processes, ensuring it was used as part of how work gets done, not alongside it.

> "It's going faster [than we expected]—both in time and in quality. The pace of the requests [for ChatGPT] keeps expanding."
>
> **Jan Andries Oldenkamp**
> CIO, Scania

04 Quality before scale

Scout24's approach highlights a quieter discipline that separated serious deployments from rushed ones: defining what "good" meant before scaling.

In building conversational real-estate search, the team invested heavily in evaluation. Custom testing frameworks, inspired by OpenAI Evals, were paired with broad internal testing to surface edge cases and calibrate trust ahead of launch.

Features were delayed when quality thresholds were not met. Trust was treated as a prerequisite, not an outcome.

The result was not slower progress, but fewer reversals — and a stronger foundation for scale.

> "At Scout24, we learned that defining 'good' before scaling AI is critical, because quality turns experimentation into something users can truly trust"
>
> **Gertrud Kolb**
> CTO, Scout24

05 Protecting judgment work

JetBrains framed AI adoption around a different productivity question. Speed matters, but the larger opportunity lies in improving the quality of expert judgment.

Developers do not just write code. They review it, reason about it, and design systems. AI was introduced to support these higher-order activities rather than replace them.

The goal was hybrid workflows that lift the ceiling on human capability while preserving the deep work that underpins long-term quality.

> “It’s not just about generating code—it has to be safe, readable, and maintainable"
>
> **Kris Kang**
> Head of Product, JetBrains

What the results signal

Across the *Frontiers of AI* stories, impact showed up through consistent operational signals:

Time returned to high-cost professionals by reducing administrative burden

Strong internal adoption in highly regulated environment

Measured efficiency gains paired with quality and satisfaction guardrails

Strong bottom-up pull and team-based capability building in operationally complex environments

Evaluation systems that make AI trustworthy in high-stakes consumer journeys

A shift from raw speed to sustained excellence by protecting deep work

The organizations pulling ahead are not simply moving faster. They are moving more deliberately.

Where leaders are heading next

The direction of travel is consistent. Leaders are moving from individual productivity gains toward workflow-level automation, and from assistance toward agents — while keeping human oversight firmly in place.

The frontier is not defined by louder claims or faster launches. It is shaped by quieter execution, grounded in evidence, trust, and scale.

Leadership lessons

01	AI adoption is a leadership discipline. Outcomes are shaped through participation, not sponsorship alone.
02	Trust is designed. Governance and evaluation make scale possible.
03	Workflow fit beats feature breadth. Adoption follows usefulness.
04	Quality unlocks speed. Defining “good” early prevents costly rework later.
05	Culture compounds momentum. Empowered teams — not just individual champions — move faster than central mandates.
06	AI amplifies people. Durable gains come from hybrid systems where humans remain accountable.

The Frontiers of AI leadership checklist

Use this checklist as a starting point to set executive accountability, align cross-functional partners, and protect quality as you scale.

Leadership & ownership

- [ ] Do we have a single executive accountable for AI outcomes, not just enablement?
- [ ] Are leaders actively using AI themselves, or only sponsoring it?
- [ ] Have we been explicit about where AI should *not* be used?

Trust & governance

- [ ] Are security, legal, and compliance partners in design, or reviewers at the end?
- [ ] Can we explain to employees why they can trust AI outputs?
- [ ] Do we know which decisions must always retain human oversight?

Adoption & workflow fit

- [ ] Are we redesigning workflows with AI, or layering AI onto existing ones?
- [ ] Do teams have permission to experiment safely without fear of getting it wrong?
- [ ] Where are we seeing the strongest organic pull — and do those teams have the mandate to turn experiments into shared capability?

The Frontiers of AI leadership checklist

Quality & scale

- [ ] Have we defined what “good” looks like for AI outputs in our context?
- [ ] Do we have robust evaluation or review mechanisms in place?
- [ ] Are we willing to slow launches when the bar is not met?

Go deeper

Watch the interviews and read the full *Frontiers of AI* stories:

- [ ] Philips — How Philips is scaling AI literacy across 70,000 employees
- [ ] BBVA — From pilot to practice: how BBVA is scaling AI across the organization
- [ ] Mirakl — Inside Mirakl’s agentic commerce vision
- [ ] Scout24 — Building the next generation of real-estate search with AI
- [ ] JetBrains — Reshaping how the world writes code
- [ ] Scania — How Scania is accelerating work with AI across its global workforce