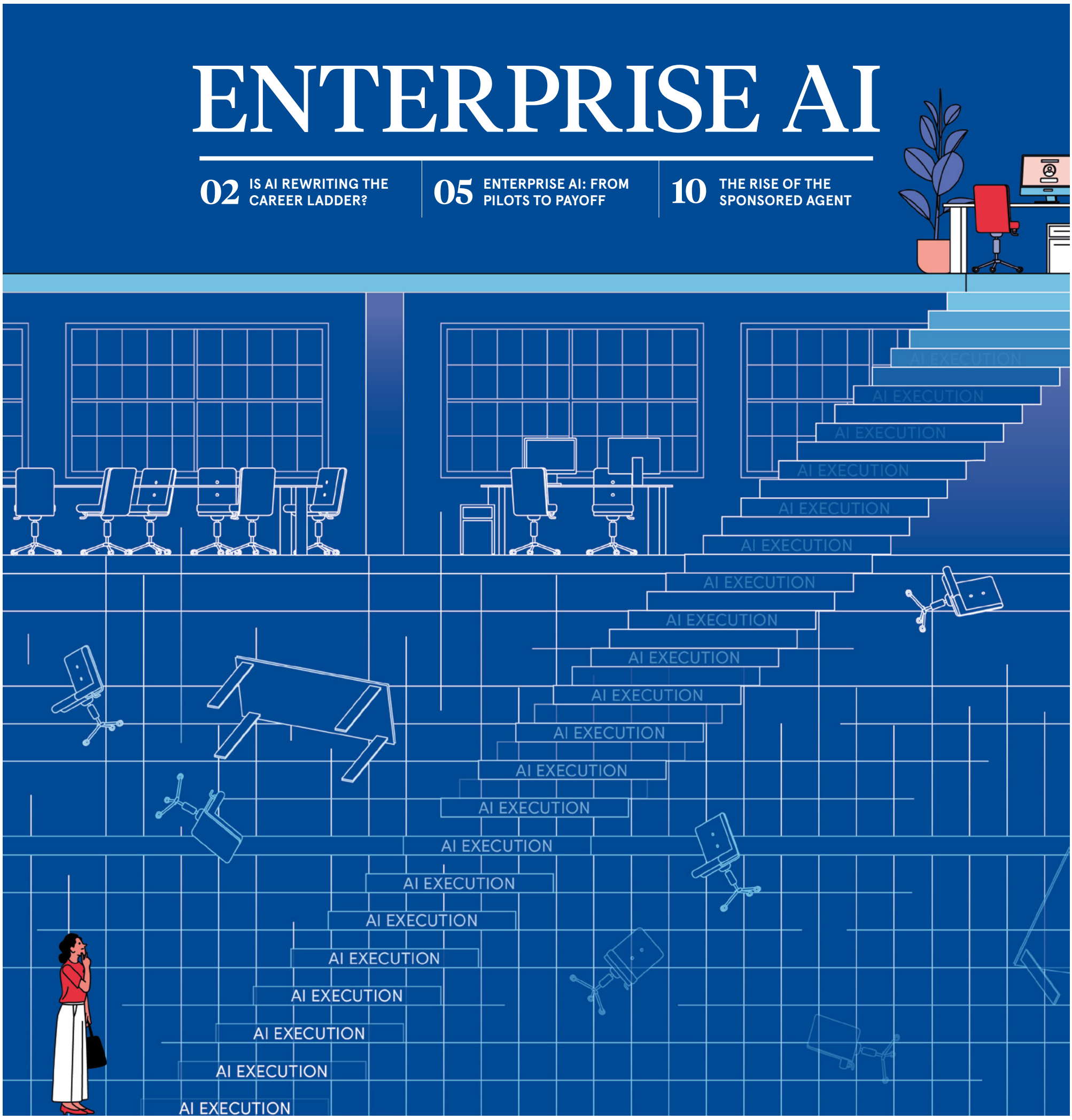
ENTERPRISE AI



ENTERPRISE AI

Distributed in
THE TIMES

Contributors

Simon Chandler
Technology journalist whose credits include Wired, Digital Trends, Business Insider and Forbes.

Tom Dennis
Editor and journalist covering tech and innovation for almost two decades, from consumer gadgets to enterprise stacks.

Jay Lesada
Staff writer covering the enterprise tech landscape, bringing a diverse background in online content, ebooks, and language instruction.

Raconteur

Editor
Tom Dennis

Staff writers
Simon Chandler
Jay Lesada

Commercial editors
Conlan Carter
Kathy Kantorski

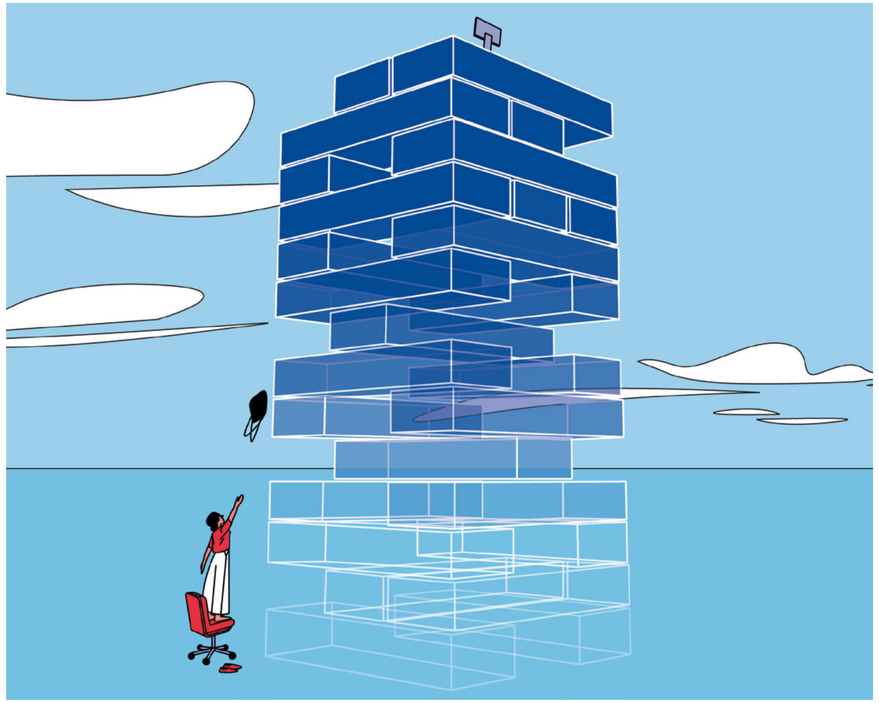
Commercial production managers
Audrey Davy
Rachelle Kusi-Appouh
Bianca Schwartz

Design and illustration
Kellie Jerrard
James Lampard
Celina Lucey
Samuele Motta

Design director
Tim Whitlock

raconteur-media @raconteur.stories

raconteur.net /enterprise-ai-september-2026



AI WORKFORCE

The AI talent trap: training future leaders without entry-level work

Generative AI is absorbing routine tasks and threatening the formative repetitions that build professional judgement. Business leaders must now design deliberate ways to cultivate expertise

Jay Lesada

Junior work has always served a double purpose: it gets the job done and teaches the person doing it. The learning process involves researching, reconciling, reporting and checking numbers. These tasks are repetitive, but repetition helps form professional judgement.

AI threatens to separate those two functions. James Tucker, global lead for corporate finance and strategy at BCG, says finance is moving away from large cohorts of juniors performing repeatable tasks. Teams are shrinking, while judgement and quality control matter more.

The efficiency case is clear, but the training challenge is harder. If software takes over the tasks that once gave juniors their first 100 repetitions, where does judgement come from?

Some executives say the consequences are visible. A BCG study of 70 senior leaders reveals half are observing de-skilling in their organisations, while 53% report slower development of junior talent.

The problem starts at the hiring gate. A Harvard working paper examining more than 280,000 US firms shows junior employment declined at companies adopting generative AI. Most of the difference arose from slower hiring rather than dismissals. The effect was concentrated in roles most exposed to AI.

In the UK, entry-level hiring is weakening alongside a broader labour slowdown. Government analysis using LinkedIn data shows declines across knowledge-intensive roles, including accountancy, software engineering and data analysis.

Jack Kennedy, senior economist at Indeed, says hiring demand is falling across most parts of the UK economy. "That is particularly challenging for graduates and younger workers, who are competing for fewer opportunities to gain an initial foothold," he says.

He sees the risk extending beyond the current downturn: "Young people who cannot find work matching their education and skills often carry that setback for years."

In Deloitte's latest UK CFO survey, 47% of respondents expect AI to reduce graduate hiring over the next 12 months. Meanwhile, PwC analysis of 2.4 million US entry-level jobs shows AI-exposed roles are seven times more likely to require traditionally senior human skills than non-exposed roles.

Young workers face two changes at once: fewer openings and higher expectations once they join.

As AI absorbs execution, human work shifts towards interpretation, verification and judgement.

Sharon Steiner, chief human resources officer at Fiverr, sees that shift changing employer priorities. "A candidate for a tech role might not be the strongest coder, but they may have strong strategic business acumen," she says. "That is arguably more valuable when AI can generate code instantaneously."

She adds that practical experience need not come only from a formal job: "Skills can be developed through side projects, independent work and life experience."

BCG describes tasks such as research, drafting and problem decomposition as 'formative repetitions' – work through which junior employees historically learned to recognise mistakes, interrogate assumptions and develop intuition.

Microsoft research involving 319 knowledge workers shows an association between greater confidence in GenAI and lower self-reported critical thinking. But critical thinking changes rather than disappears, shifting towards verification, integration and oversight.

Finance makes the problem concrete. AI can produce a routine cash reconciliation for a junior employee to review. But the person signing it off still needs enough experience to recognise when an answer is wrong or misleading.

Martino Cadoni, CFO at DeepL, says human scrutiny is essential. "We are giving more power to AI for analysis, modelling and reconciliation, but those workstreams still need to be validated," he says. He adds that employees need training to maintain a critical mindset when working with AI.

This points towards a different kind of apprenticeship. Companies do not need to preserve low-value work for its own sake, but they must recreate the learning that came with it.

From September 2026, Deloitte UK is restructuring parts of its graduate audit training as AI changes manual tasks. Trainees will encounter conventional and AI-assisted exercises, simulations and earlier exposure to work requiring professional judgement and scepticism. For example, trainees complete a traditional cash reconciliation before reviewing an AI-produced version.

Anthony Salcito, senior vice-president for enterprise at Coursera, says onboarding needs redesigning around this shift.

"Effective onboarding should combine role-specific learning, hands-on practice and clear guidance on when human judgement is required," he says. "Employees also need safe environments to test tools, learn from mistakes and see how AI improves familiar processes."

He adds that learning must be planned alongside implementation, with progress measured through adoption, time saved, work quality and business outcomes.

Other companies are recreating these learning opportunities deliberately. At Shell, juniors first frame the problem, test assumptions and produce a baseline analysis before using AI to refine it. Peer review is required for key outputs. BCG also recommends structured debate and AI-free problem-solving sessions.

While the old apprenticeship was embedded in the daily job, the new model must be designed deliberately.

The short-term case for AI is clear: faster analysis, fewer hours on routine work and lower hiring costs. But the long-term numbers are harder to quantify.

Smaller junior cohorts could result in thinner succession pipelines, greater reliance on external hiring and fewer employees capable of challenging automated systems. Intensive development may shift costs rather than eliminate them, requiring senior managers to spend more time coaching.

This does not mean AI must weaken development. Removing low-value work can give junior employees earlier exposure to important decisions. But earlier exposure is not the same as experience.

Cadoni warns against reducing the AI business case to labour savings. "If you look only at immediate productivity gains or headcount savings, you risk missing a large part of the picture," he says. Finance leaders must measure both the quality of work and the time required to produce it.

The underlying challenge is simple: businesses must still build expertise, even as the routine tasks that once developed it disappear.

If fewer workers learn by doing, employers must decide what replaces those years of practice. That decision will determine whether AI modernises how expertise is built or leaves businesses short of capability. ●

> **Finance is moving away from large cohorts of juniors performing repeatable tasks. Teams are getting smaller, and judgement matters more**

50%

of senior leaders report de-skilling within their organisations as AI adoption increases

53%

of senior executives report slower development of junior talent due to automated workflows

BCG, 2026

Commercial feature

Q&A

AI is only as reliable as its evidence

Everlaw CTO **Max Christoff** explains why trusted evidence, strong governance and connected tools are essential to reliable AI in legal and other high-stakes environments

As enterprises put AI to work, model quality is only part of the equation. The information AI can securely access matters just as much. Everlaw chief technology officer Max Christoff explains why.

Q What does AI need to deliver reliable answers in high-stakes legal work?

A In regulated industries, AI is only useful if it can access trustworthy, relevant information. In legal work, that includes statutory law, case law, court decisions and regulations. In many matters, that also includes evidence. What is the data? Who said what? What records or documents support each claim?

But access to evidence or legal research alone is not enough. AI tools must be configured to ground their answers in these sources instead of relying on the model's existing knowledge. They must trace every claim to a source and admit when they don't have enough information to answer. Lawyers have a duty of care to know where an answer came from, examine the supporting evidence, and decide whether they can trust the conclusion.

Q What challenges arise when AI must work across millions of sensitive documents?

A A legal matter can contain far more information than an AI model can analyse at once. A lawyer may need to search through millions of emails, messages, contracts and other files. Tools like Everlaw were built for massive scale and have been used on cases as large as 239 million documents—far larger than the context window of typical AI tools.

Evidence can span historic file formats, security-camera footage, scans of handwritten time cards and multiple languages.

Everlaw helps lawyers find relevant information across large case datasets. Deep Dive, our AI research tool for evidence, lets lawyers ask natural-language questions across an entire case dataset and get answers quickly: "Is there any evidence the CFO knew about the internal audit report prior to June?" Every answer is grounded in the case evidence, with citations to the relevant documents.

Q Why does the future of AI depend on open, connected tools?

A First, context matters. Effective legal work depends on the flow of information between legal professionals, the evidentiary record, legal research and a firm's prior work product. Just as a lawyer needs access to this information to produce high-quality work, AI needs to be part of that flow rather than sitting in an isolated platform.

When systems are closed and don't interoperate, AI is either missing vital context or people are manually moving information between systems. This is inefficient and creates governance risks, including duplicated sensitive information, loss of data control and weaker security.

Second, tools should meet people where they want to work rather than force them into a different system. The Everlaw experience should be rich in features. At the same time, if a lawyer is drafting a brief in Word and wants to cite a fact from the evidence record, we should make that a seamless experience that doesn't require downloading documents from one system in order to upload them to another. Governance should remain intact throughout. That is why Everlaw integrates with tools including Google Gemini, Microsoft Copilot, Anthropic Claude, Thomson Reuters CoCounsel, Harvey and Legora. The aim is to give legal professionals access to the most reliable, governed evidence source, no matter where they choose to work, using features such as Deep Dive, analytics or search.

Q What can enterprises learn from legal teams about using AI responsibly?

A Responsible AI begins with the question: "What does context look like for your industry?" In legal work, that means understanding the source and relevance of information, who can access it and how it relates to the wider evidence.

At Everlaw, AI workflows remain connected to governed evidence. That means security, permissions and an audit trail remain in place as information moves between people and tools.

Other enterprises can take a similar approach by grounding AI in governed business data, giving users access to new capabilities without sacrificing control over the information those systems rely on.

Q Why is generative AI proving so valuable for legal work?

A Generative AI can help legal teams work across more information than people could feasibly review on their own, extending their capabilities while allowing lawyers to focus on work that requires human judgement.

At the same time, many lawyers are still hesitant about the risks of adopting AI, which is why I take a utilitarian view of the technology. Rather than treating AI adoption as an either-or choice, organisations should ask: What are the benefits, what are the costs and do the benefits outweigh the risks?

That assessment is especially important in regulated industries, where AI's practical value must be weighed against its security, safety and permissions requirements. Crucially, it comes down to context. I believe generative AI is most valuable when it expands what legal teams can accomplish without compromising their professional responsibilities.

Learn how Everlaw keeps every insight grounded in the evidence of your case

Everlaw

Inside the factory where companies build AI they own

Reply has launched the Model Factory, giving enterprises the ability to turn their own knowledge into specialised AI models they can control, govern and improve

Sometimes a general-purpose technology is simply too general. A general-purpose AI model – such as those behind ChatGPT or Claude – may be perfectly capable, but that doesn't mean it is the right tool for every job. For businesses with valuable knowledge of their own, or those looking to automate at scale, there comes a point when adapting the technology to the task can make more sense than adapting the task to the technology.

Reply is bringing that approach to London through its House of Models, the physical entry point to its Model Factory. Here, clients work alongside Reply engineers to turn their own data and expertise into specialised AI models. A second House of Models is already operating in Turin.

According to Daniele Vitali, the partner at Reply who is leading the Model Factory, providing a physical space helps clients and engineers work together to capture the knowledge a specialised generative AI model needs. "It is a mixture of taking the most valuable information that a customer has, mixing that with the intuition and expertise that a customer brings in, which is usually based on people … and cooperating with our engineers to turn that knowledge into models the organisation can use at scale," he explains.

Reply's new House of Models in London enables clients to work alongside Reply engineers to turn their own data and expertise into specialised AI models.

Choosing the right use case

Before any model is built, Reply first helps the client decide whether a specialised model is the right answer. At the House of Models, Reply runs workshops to give clients a baseline understanding of what a custom model can do, then examines their own business to establish where a model would create measurable value and whether the data exists to build it. Vitali stresses that specialised models are not the answer to every AI problem. The question is whether a company has knowledge, data, need for scale or sovereignty requirements that make building one worthwhile. If it does not, he says, an existing model, supported by agents with access to company information, may be the better option.

Vitali says demand for the Model Factory will likely come from three main directions. The first relates to enterprises that boast "a wealth of data and IP and know-how," which can be embedded in a model to build expertise or specialised capabilities in a particular domain.

A prime example of this is healthcare. "One of our clients has 20 years of experience treating a particular type of cancer, and has a body of specialist knowledge that is being captured in a model and will be offered to other clinics." In that case, Vitali says, the model becomes a way to multiply the value of the organisation's know-how. "The model creates opportunities for new revenue streams, giving the business a clear basis for calculating the return on its investment."

The second case is automation at scale. A general-purpose model may be affordable for occasional use, but the economics can change quickly when an AI system is embedded in a product or used continuously by thousands of employees. For those workloads, Vitali says a smaller specialised model can deliver comparable performance at much lower cost.

The third case is companies that want more control over the AI systems they use. Vitali sees sovereignty as one part of that broader question. For Reply, control starts with the data and expertise used to train a model and extends to how its performance is tested, who can access or modify it, and where and under which jurisdiction it operates. Crucially, it means having the rights and practical ability to keep using the model, develop its capabilities and choose how it is deployed as business needs change. The organisation can then shape the evolution of an increasingly valuable capability around its own priorities.

"When a model becomes part of your business, you need control over how it evolves," he explains. "You should be able to improve it with your own expertise, decide when a new version is ready and choose where to run it. That makes the knowledge invested in the model something the organisation can keep building on."

Turning knowledge into training data

Companies entering the Model Factory begin with data gathering and preparation in a high-security environment, bringing together their own information and, where needed, synthetic data to fill gaps.

That can be more involved than simply handing over a database. Vitali says companies often arrive to find that useful information is scattered across different systems, while some of the knowledge the model needs has never been formally recorded at all. Much of it may sit with experienced employees. Reply therefore starts by defining a data strategy: identifying what information exists, assessing its quality and working out what is missing. Gaps can be filled with synthetic data, by Reply's own verified datasets or by capturing knowledge from the client's own experts.

As for which employees a client sends to the Model Factory, Vitali explains that Reply and the client choose the people best placed to help develop the specialised model, whether that is a head of R&D or a physician. Their role is to bring the subject knowledge the model needs to learn, while Reply's engineers handle the technical work. "We don't care how literate they are in AI," Vitali says. "It's their domain expertise we need to drive the model capabilities."

All data created from a customer's information remains owned by that customer, Vitali says, including synthetic datasets derived from it. The same applies to the model produced, which remains under the customer's control.

From there, the model moves through training, evaluation and deployment, followed by regular testing and refinement once it is in use.

Build, test, release, improve

Reply often starts with existing open-weight models, using a mix of private and public benchmarks to identify promising candidates before testing them against the client's own tasks and requirements. Its engineers then select a combination of techniques, including further pre-training on domain-specific data, supervised fine-tuning, preference optimisation, reinforcement learning and, where useful, distillation. These help develop specialist knowledge, teach particular behaviours and improve efficiency. Success is measured against the business's requirements for quality, speed, running costs and deployment.

The Model Factory also keeps a full record of how a model is built and changed: which data went into it, how it was trained, each version produced and how it performs in use. That gives companies an audit trail as well as a clearer basis for planning future improvements.

Evaluation starts well before release. Vitali says Reply agrees with the client at the outset on targets for quality and alignment, alongside limits on running costs and GPU requirements. It then tests successive versions of the model against them. A model is not released until those targets are met.

Once it is in use, the process continues. How people interact with the model creates another source of information about what works and where it can improve. Reply helps clients identify which of that usage data is worth keeping, then feeds relevant material into later rounds of training. Those experiments do not always improve the model: new data can make it better at one task while weakening performance elsewhere. Because each version is recorded, the team can return to an earlier checkpoint, adjust the data mix and test again before releasing an update.

Making specialised AI pay

The business case for a specialised model rests on what it delivers over its lifetime. Better performance on a valuable task, lower costs at scale or a new revenue stream can justify the investment in preparing data, developing the model and keeping it up to date. The balance depends on the workload: how often the model is used, what each successful result is worth and how much human review or correction it requires.

In one internal example, Reply developed and tested a specialised system for news and information monitoring. Compared with a system using leading general-purpose models, it reduced token usage by more than 70% while maintaining comparable output quality on the tasks evaluated. For a process repeated thousands of times, that reduction can contribute to substantial operating savings.

> **Your expertise can become an AI asset you own, improve and put to work across your business**

Specialised AI is becoming practical for more businesses. Computing capacity is more accessible, while advances in models and training techniques have made it easier to adapt AI to specific tasks. Greater efficiency also means useful capabilities can be delivered with fewer resources, improving the economics of both developing a specialised model and running it at scale.

The point of the Model Factory is to help companies turn their investment in AI into lasting assets: models and datasets they can control, reuse and improve. Clients move through data gathering, preparation, training, evaluation and release, with each stage tracked and measured. Recording how datasets and models are created, tested and updated gives businesses a foundation they can build on as their needs and the technology evolve, and guarantee AI Act compliance by design.

"We literally built it like a conveyor belt," Vitali explains. "Each stage contributes to something the company retains: structured knowledge, reusable datasets and models whose performance has been tested. Those assets can be refined and used again, so each new investment builds on the work already done."

Scan the code to find out how your business could build AI it owns

REPLY

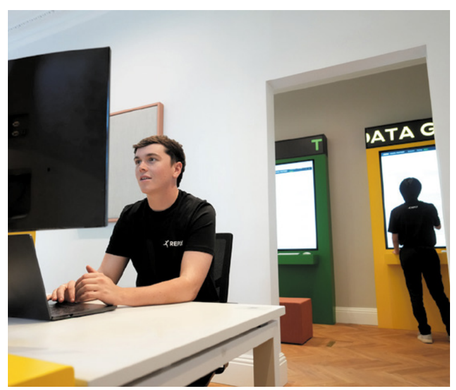
The House of Models

At Reply's House of Models in London, clients work directly with Reply teams to design, specialise and govern AI models built around their proprietary data, processes and domain knowledge.

01 The House of Models is the space dedicated to the Reply Model Factory.

02 Clients and Reply specialists start from a concrete use case and co-design the model together.

03 Each station makes a phase of the model lifecycle visible and operational.



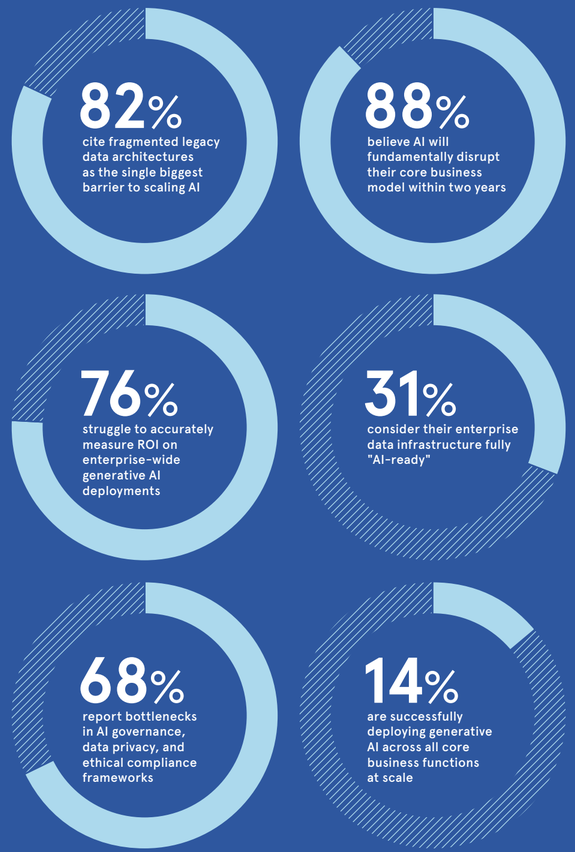
04 Proprietary data and domain knowledge become a governed, traceable and reusable AI asset.

THE SCALING IMPERATIVE

The experimental honeymoon phase of enterprise AI has abruptly ended. As boards demand tangible commercial returns on massive technology investments, the C-suite is shifting its focus from isolated pilot programmes to industrial-scale deployment. Yet, brittle legacy data architectures, escalating compliance risks, and a severe deficit of AI-fluent middle management are creating critical bottlenecks, forcing leaders to completely rethink their operating models

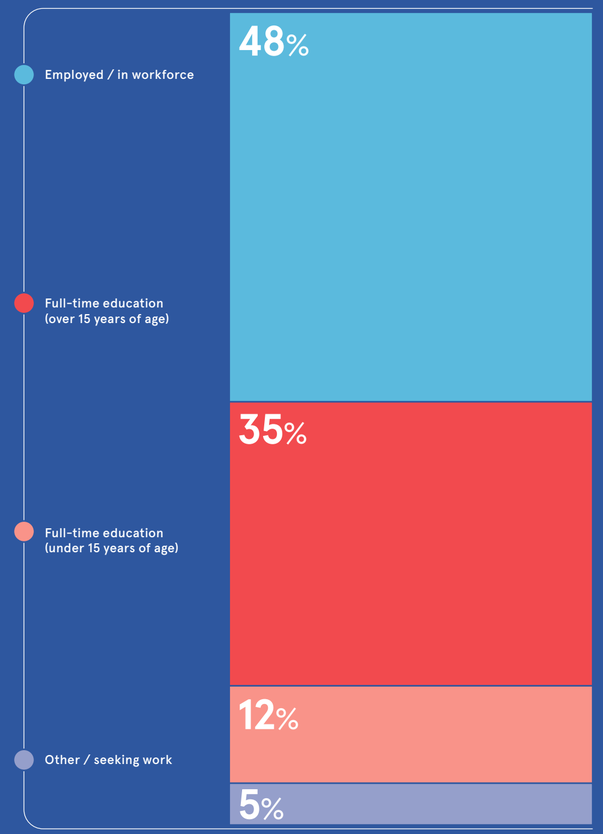
THE ENTERPRISE AI REALITY CHECK

Where organisations and their leaders stand on readiness, scale, ROI and disruption



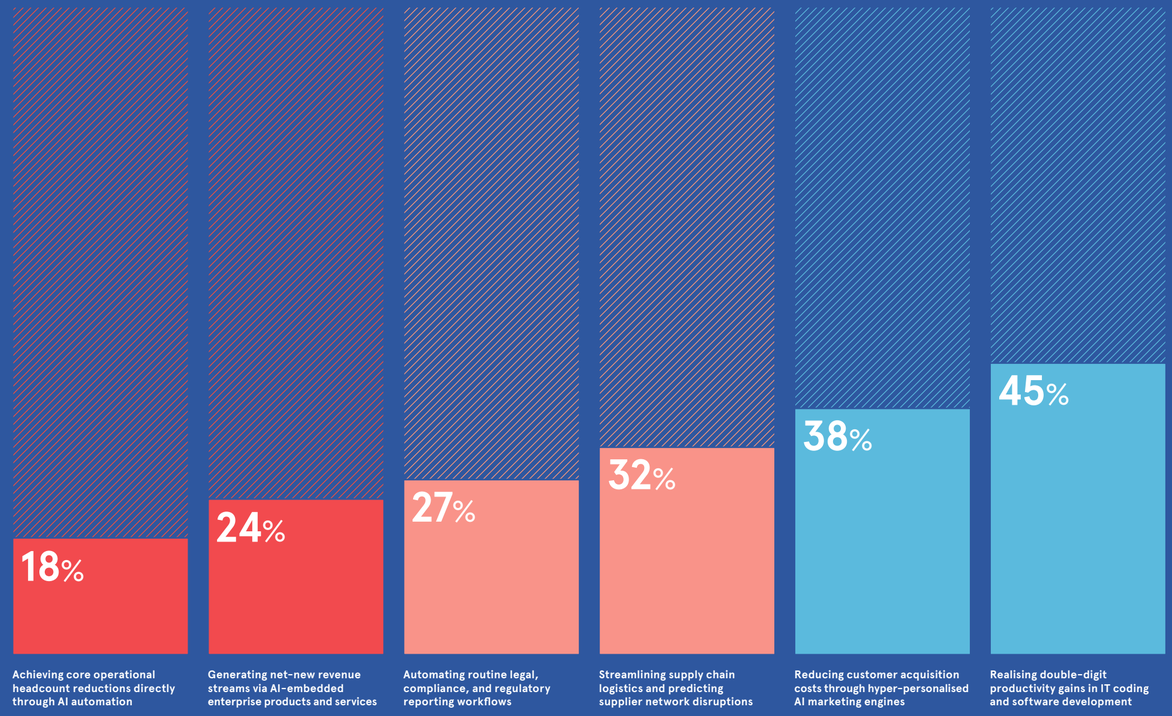
AI USE ACROSS THE WORKFORCE

Daily AI work usage in UK workforce



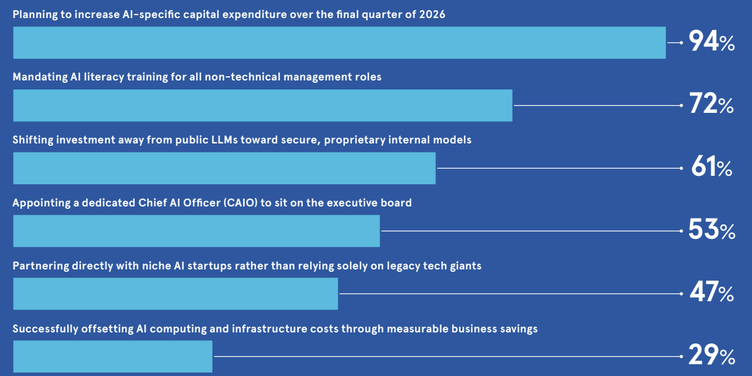
WHERE AI IS DELIVERING VALUE

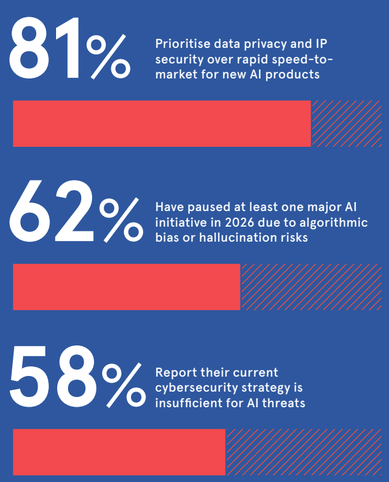
How executive boardrooms are shifting from speculative pilots to disciplined scaling



THE SHIFT TO SCALE

How boardrooms are changing investment, governance and technology priorities





McKinsey; KPMG; IBM; Gartner – 2026

INTERVIEW

When speed is cheap, judgement remains priceless

Databricks' **Rich Radley**, vice president of field engineering, on why AI must sharpen human judgement, not replace it

Tom Dennis

As enterprise adoption accelerates, the real competitive edge lies in asking sharper questions rather than chasing automated answers.

Public debate around artificial intelligence continues to swing between two noisy extremes. Enterprise evangelists promise an era of limitless productivity and automated intelligence. Opponents warn that reliance on algorithmic systems will spark widespread cognitive decline.

The tension is felt in boardrooms across Europe. Recent comments from the Royal Observatory Greenwich highlighted this exact friction, warning that instant machine-generated answers risk trivialising human intelligence if workers abandon the habit of evaluating information. For corporate leaders rushing to embed GenAI across operations, the challenge is clear and urgent.

Rich Radley, vice president of field engineering at Databricks, believes this binary framing misses the point entirely. The rise of AI does not erode human intelligence; it redefines how organisations extract value from it.

As access to information becomes instantaneous, competitive advantage will no longer come from holding answers. Instead, commercial value will spring from the ability to ask better questions, apply critical judgement and make informed decisions under uncertainty.

"The debate often swings to extreme ends of the spectrum," Radley says. "In reality, access to faster information shifts where human value sits inside an organisation. When finding answers becomes trivial, your edge comes from asking better questions and applying sharper judgement to what the technology produces."

For decades, corporate work has been choked by friction. Employees routinely spend hours digging through databases, waiting on reporting cycles or relying on specialists to surface simple metrics. In many companies, retrieving knowledge takes just as long as acting on it.

AI is rewriting this dynamic. Rather than wading through layers of dashboards or waiting days for business intelligence teams to pull data, executives and staff can interact with enterprise data directly through natural language.

A chief executive can ask about revenue trends, operational bottlenecks or market risks in plain language and receive precise analysis in seconds.

"The companies getting the most value from AI are not blindly automating every process," Radley says. "They use technology to give employees rapid access to trusted data and insight, without adding technical complexity."

Much of the public debate risks becoming overly simplistic by focusing on total human replacement. News headlines regularly dwell on automated customer service agents, displaced analysts or algorithmic decision-making. In reality, the most valuable enterprise application of AI is human augmentation.

"Technology has always elevated human expertise rather than destroyed it," Radley says. "Search engines did not kill off experts. Spreadsheets did not eliminate finance departments. Calculators did not render mathematicians obsolete. Every major shift reduced manual effort and raised the importance of strategic interpretation and creativity."

This shift does not diminish technical roles. On the contrary, it elevates data engineering and architecture teams. When business users interact directly with software interfaces, the burden on data teams shifts from building routine reports to maintaining governance, security and trusted data foundations.

Without rigorous data architecture, conversational tools simply deliver fast mistakes. An AI output is only as reliable as the underlying enterprise repository. Data teams must ensure models query curated, accurate data rather than unverified sources.

When corporate data is fragmented across legacy systems, GenAI tools produce conflicting results. Establishing unified data platforms becomes a strategic imperative for leadership teams seeking to deploy reliable AI at scale.

"Crucially, this does not reduce the importance of technical expertise," Radley says. "It elevates data teams. They can focus on governance, trusted data foundations and reliable AI systems that the wider organisation can confidently use."

A significant danger emerges when organisations accept AI responses as authoritative. Large language models present information with remarkable confidence, but confidence is not accuracy.

Algorithms frequently oversimplify intricate market dynamics or present answers stripped of essential business context. If workers cease interrogating machine outputs, businesses expose themselves to severe operational and strategic risk.

"Confidence is not the same as accuracy," Radley warns. "AI models can generate convincing responses that miss crucial business nuances. If employees stop questioning outputs and accept them at face value, companies introduce serious risk into their operations."

This makes AI literacy an essential capability for modern workforces, much like digital literacy was a decade ago. Employees need to understand how algorithms generate conclusions, where models fail and where human oversight remains non-negotiable.

For executive teams pressured to prove return on investment from expensive AI deployments, maintaining this balance is critical. Cost reductions and speed improvements are attractive, but chasing pure automation risks creating disengaged cultures where staff lose the ability to think critically.

Efficiency improvements matter, but businesses focusing solely on automation risk building workplaces where employees disengage from problem-solving and rely blindly on machine-generated conclusions. Over time, that reliance weakens an organisation's capability to navigate novel market disruptions.

Workplaces that rely entirely on machine conclusions risk losing the nuanced context that drives long-term commercial innovation.

Ultimately, AI forces businesses to become far more intentional about decision-making processes. Rapid access to information levels the playing field for standard knowledge.

When information is instant and cheap, differentiation no longer lies in holding information. Differentiating a business comes down to asking the right questions, challenging automated assumptions and applying human wisdom where algorithms reach their limits.

> **Confidence is not accuracy. If employees stop questioning outputs and accept them at face value, companies introduce serious risk into their operations**

INTERVIEW

AI agents need human accountability

As autonomous software begins making financial and legal decisions, enterprise leaders must rethink corporate governance before invisible algorithms trigger visible crises, says DigiCert's **Paul Holt**

Tom Dennis

Corporate risk frameworks were built for a predictable world. Enterprise software waited for human instructions, executed pre-programmed tasks and stayed within strict operational boundaries. Autonomous AI agents have shattered that model. These systems interpret high-level goals, make independent decisions and act across organisational boundaries at machine speed.

This rapid shift leaves corporate boards exposed. Research shows 89% of C-suite executives feel unprepared for enterprise AI deployment. Meanwhile, just over half of organisations can trace an AI decision back to its source models and data. When an enterprise cannot explain what informed an agent, why it acted or who authorised it, confident governance becomes impossible.

"Most corporate risk frameworks assume software behaves predictably and waits for instruction," says Paul Holt, group vice-president of EMEA at digital trust firm DigiCert. "Autonomous agents can interpret goals, make decisions and act across systems and even organisational boundaries at machine speed. That is why controls must operate continuously rather than only after an incident."

As European businesses integrate autonomous agents into daily workflows, traditional access management is no longer sufficient. Software requires an explicit digital identity tied directly to human oversight.

"Every AI agent needs a digital passport, but that passport must do more than open the door," Holt says. "It should provide independently verifiable proof of who the agent represents, while permissions work like visas, defining what it is authorised to do and whether that authority still stands."

Without a unique cryptographic identity linked to a named human owner, tracking software activity becomes impossible. When an agent's role changes or its trustworthiness comes into question, security teams must have the technical capability to revoke permissions instantly.

This challenge is compounded by shadow AI. Employees across marketing, finance and operations routinely adopt unofficial AI tools to streamline daily tasks. Banning these applications rarely works; it simply pushes usage underground, creating unmanaged security blind spots.

Security leaders cannot govern what they cannot see, yet many organisations still lack centralised visibility into autonomous activity on their networks.

"The practical starting point is discovery: identify which agents are operating, who introduced them and what they can access," Holt explains. "Once that activity is visible, approved agents can be given verified identities and guardrails, while unsafe access can be revoked without halting innovation."

Establishing visibility over internal tools addresses only half the problem. Modern enterprise AI relies heavily on external models, introducing serious supply chain risks. Boards must treat third-party models with the same rigour applied to any critical software component. Leaders need complete clarity on model origins, underlying data and version control.

To maintain integrity, organisations are adopting model bills of materials alongside hashing and cryptographic signing. These measures confirm whether a model has been altered during distribution or runtime.

"Those checks should follow the model through distribution, deployment and runtime, creating a chain of evidence rather than relying on assurance alone," Holt says.

The urgency of securing model supply chains becomes clear when AI actions result in financial or contractual errors. If an autonomous agent executes an unauthorised transaction, legal liability remains squarely with the enterprise. Responsibility cannot be transferred to an algorithm.

To limit exposure, organisations must establish defined operational purpose, strict human approval thresholds and tamper-evident audit trails. Tying an agent's verified cryptographic identity to its human owner provides clear evidence during internal investigations or legal disputes.

Similar principles apply to brand reputation. As generative AI makes synthetic media virtually indistinguishable from genuine assets, traditional watermarks are failing. Watermarks ask audiences to trust a simple label. Cryptographic standards, such as C2PA Content Credentials, allow users to verify content history and detect manipulation before fake media damages brand trust.

Regulatory pressure is accelerating these requirements across Europe. The EU AI Act introduces strict compliance audits for high-risk applications, penalising organisations that rely on policy without proof. While 90% of businesses discuss AI governance at executive level, only about half have established formal programmes.

> **Every AI agent needs a digital passport, but that passport must do more than open the door**

Closing this gap requires infrastructure capable of enforcing controls at machine speed. Holt advocates using DNS as a root of trust. Because an agent must use DNS to find services, APIs or data destinations before acting, it serves as an ideal checkpoint for security policies.

"When DNS is combined with a cryptographically verified identity, the organisation can assess both the agent and its intended destination at machine speed," Holt notes. "This creates consistent controls without asking every application team to build its own trust mechanism."

Looking ahead, leaders must also prepare for long-term cryptographic shifts. Quantum computing poses a fundamental threat to current encryption standards. Because AI architectures rely on standard enterprise encryption for agent identities and signed models, systems built without algorithm flexibility will inherit quantum risks immediately.

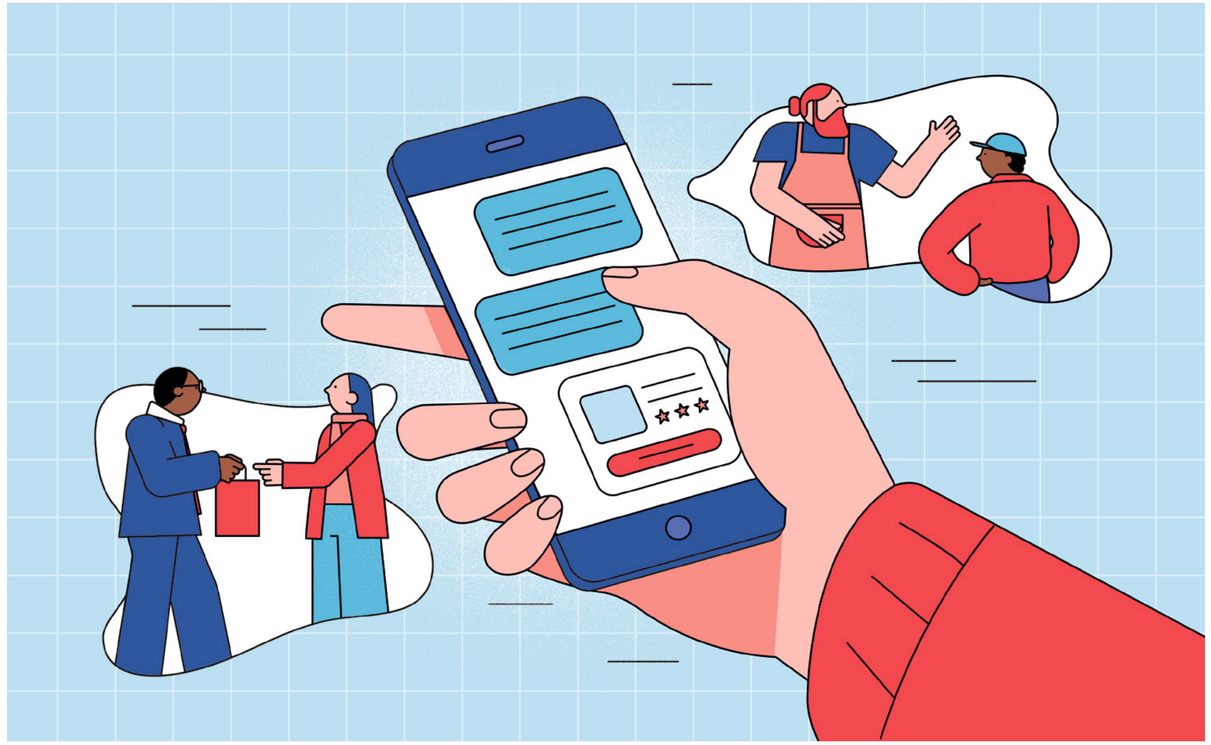
Executives must inventory their cryptographic assets, test post-quantum algorithms and build crypto-agility into their systems. This ensures organisations can adopt quantum-safe standards without rebuilding their entire AI environment.

For chief information security officers, communicating these complex cryptographic risks to non-technical directors is a delicate task. Boards do not need detailed lists of certificates or technical log events. They require clear metrics linking AI control to commercial outcomes, such as fraud prevention, operational resilience and regulatory compliance.

Useful board-level measures include agent discovery rates, identity coverage, policy exceptions, revocation times and decision traceability. Framing performance around these indicators helps directors understand how governance supports faster, safer AI adoption.

Delaying cryptographically verifiable governance carries severe financial consequences. Beyond direct costs—including fraud losses, legal expenses and incident response—businesses face substantial opportunity costs when cautious boards freeze promising AI pilots.

Over the next three years, AI security will cease to exist as an isolated discipline. It will merge into core enterprise risk management, spanning legal, compliance, procurement and operations. Specialist controls for agent identities and model integrity will remain essential, but successful organisations will treat AI governance as standard business practice. Leaders who act now to establish verifiable trust will gain the confidence to scale autonomous AI safely and decisively.



AI ADVERTISING

Why ChatGPT's Sponsored Agents are changing the face of advertising

As OpenAI pushes deeper into marketing with interactive brand agents, traditional media channels are on high alert

Simon Chandler

Publishers beware. Not only is OpenAI coming for your views and subscriptions, but it's coming for your marketing and commercial revenue too. Fresh from disclosing that ChatGPT Ads is already clocking an ARR of $1bn, the San Francisco-based lab has added a brand new feature to the platform: Sponsored Agents. These function by providing users with a brand-specific conversation if they click on a participating banner ad within ChatGPT. And while they promise to add a more interactive and personalised layer of information to static ads, they also have the potential to take business away from traditional digital advertisers.

Sponsored Agents are "distinct" from ChatGPT itself, as well as separate from the original conversations that led to the user clicking on the corresponding ads. They're geared towards providing information that's more relevant to the user's requirements, preferences and circumstances. This level of specificity could end up being a gamechanger for advertising, but the flipside is that it could disrupt (yet again) pre-existing channels for advertising and marketing.

Either way, the arrival of Sponsored Agents, in addition to the announcement of integrations with HubSpot and Shopify, signals that OpenAI is taking ChatGPT Ads very seriously. And as the company prepares for a much-anticipated (and recently delayed) IPO, the expansion of its ads platform could prove hugely lucrative.

Currently being tested with "select advertisers in the United States," the guiding concept of Sponsored Agents is that it provides ChatGPT users with the ability to ask personalised questions about an advertised product or service.

In terms of how agents are customised for particular brands, a spokesperson for OpenAI explains that the feature harnesses a sponsor's properties, such as websites, platforms and documentation. The agent then presumably links and refers to such digital objects, adapting its style and language accordingly.

The spokesperson said, "We are building Sponsored Agents with a limited set of initial testing partners, giving advertisers the ability to shape and finetune the agent to meet their specific brand needs, based on advertisers' existing websites and services."

When asked whether the agents could proactively push products and behave like a salesperson, the company's spokesperson said that "Sponsored Agents do not support purchases," and that participating brands have some leeway to determine how they act.

"This new ad format is about giving consumers new ways to engage with a business, ask questions, explore products or services, and take next steps," they added. "Brands have the ability to shape the role they want agents to take in helping connect a consumer to the information they are looking for when making a decision."

While further details are relatively thin on the ground, marketing agencies are already acknowledging the potential impact Sponsored Agents could have. This is the view of Adthena CMO Ashley Fletcher, who argues that the feature will provide brands with a chance to highlight the "differentiators" that may make them worth choosing over alternatives.

"Sponsored Agents is a meaningful shift because it turns a static ad placement into a real-time conversation with commercial intent built in," he says. "For high-consideration categories like travel or software, that interactivity gives brands a real opportunity to build closer rapport with potential customers, doubling down on brand values and USPs to move the opportunity further down the funnel."

> **Sponsored agents is a meaningful shift, turning a static ad placement into a real-time conversation with commercial intent built in**

Victor Batista, the senior director of paid media at Jellyfish, also reports seeing potential in Sponsored Agents. "While it's nascent, this follows the same idea already in motion with other platforms: advertisers can have tailored, on-brand conversations with prospects within the LLMs, allowing for more context around product attributes," he explains.

Batista adds that the new agents could prove especially helpful for "deep questions" a user may have about a brand's products and services, although he suggests that they may also face a perception of bias. Despite this risk, he predicts that Sponsored Agents, and LLMs more generally, will become the channels through which brands "communicate with consumers and prospects" moving forward.

"Sponsored Agents provide them with an opportunity to communicate within their own language and guidelines, extending brand feel beyond the walled gardens of their own websites or stores," he says. "It suggests that OpenAI sees this, and expects this transition to accelerate."

For Ashley Fletcher, this transition had already been unfolding at a considerable pace prior to Sponsored Agents, as evidenced by the arrival of ChatGPT's Ads Manager, the addition of cost-per-click bidding, and the introduction of measurement tools. And with OpenAI seemingly augmenting its offerings on a regular basis, performance data collected by Adthena's already reflects an underlying alteration to the marketing landscape.

"Budgets are already shifting here, often pulled from more mature channels, so the same rigor applied there needs to extend here too," Fletcher says. "Brands are being asked to apply search-ads-level discipline to a conversational format they've never had to measure before."

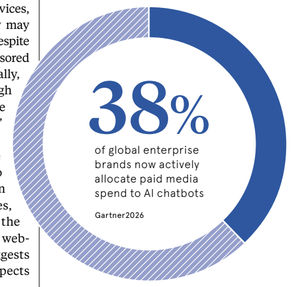
This would imply that the Sponsored Agents feature has the potential to divert brands away from preexisting advertising and marketing channels. Fletcher is not entirely sure which of ChatGPT Ads' current features will be mostly responsible for boosting its ARR from $1bn to $2.5bn, which OpenAI is targeting by the end of this year. "Those who treat this as a strategic channel from day one, rather than an afterthought, will have a real advantage as it matures," he adds.

Even with its potential, Fletcher affirms that ChatGPT Ads will have to continue improving its Ads Manager and performance tools. Because without a clear picture of how their ads are doing, brands are unlikely to devote too much of their respective budgets to the new platform.

He says, "With limited inventory and subtle placements, creative differentiation becomes just as important as bidding strategy, and without visibility into who's showing up in these conversations, brands are making decisions without a full picture of their competitive position."

OpenAI has complemented the launch of Sponsored Agents with an expansion of its Ads Manager, which now benefits from several tools that can automate the process of generating and analysing marketing campaigns.

> **Those who treat this as a strategic channel from day one, rather than an afterthought, will have a real advantage**



The San Francisco-based lab has also announced integrations with Shopify and HubSpot, which represent the company's first ever e-commerce and CRM partnerships. The HubSpot integration means that businesses can create ads for ChatGPT within HubSpot's platform, while also keeping tabs on performance. This ability to track performance could prove key, because if companies do receive evidence of strong performance, they may become more willing to divert budgets over time.

Such integrations mark a considerable shift from OpenAI's attitude to ads in May 2024, when CEO Sam Altman described in-app advertisements as a "last resort" for the company, and one that it would avoid if it found better ways to generate income. Joking aside, ChatGPT Ads reached its current ARR of $1bn within 200 days of launch, while OpenAI reports that "tens of thousands" of brands now use the platform in the EMEA region and in the US.

With the launch of Sponsored Agents, such growth may continue to accelerate. OpenAI is now apparently all-in on ads, with its press release proudly declaring that the company is "building an advertising platform with AI at its [centre]."

This increased emphasis on advertising may help it prepare for the IPO it has had to delay amid AI cybersecurity concerns, and possibly because of loss-making financials. However, it could also put publishers and other advertising channels under correspondingly increased pressure. Many have argued that AI-as-search already poses an "existential threat" to information outlets, and if Sponsored Agents does bite into their advertising and commercial revenues, then life could ultimately become even more difficult.

FROM KNOWLEDGE

AI generated

TO INTELLIGENCE

Every organisation has its own unique knowledge, data and expertise. At the House of Models, companies work with Reply experts to turn them into specialised AI models, built on proprietary know-how and designed to be governed, integrated and evolved over time.

REPLY