Where are all the AI agents?

Despite the fanfare, agents are keeping a low profile

John Leonard

16 July 2025 • 6 min read

▶ Listen to this article
7 min

SHARE

Image: Where are all the AI agents?

It struck me recently that, for all the hype, in the 18 months since I first wrote about AI agents I've yet to see one 'in the wild'.

I've seen plenty of tech company demos where agents whizz around 10X-ing this and hyper-personalising that, and I've listened to presentations by execs who have apparently turned their back offices over to autonomous actors, but on further digging these have all turned out to be proofs-of-concept or rebadgings of RPA to fit with the zeitgeist.

It's hardly unknown for tech hype to run ahead of reality, of course, but given all the razzamatazz, these agents really are surprisingly incognito. Here are a few thoughts as to why that might be.

The hype is ahead of itself

After OpenAI astonished the world by releasing ChatGPT, competitors realised they had some serious catching up to do. The concept of autonomous software agents has been around for a while, add the magic word AI and bingo.

Which is not to dismiss agentic AI as vapourware. It is a logical progression after all, it's just not as mature as the marketing would have us believe.

There are numerous agents and platforms out there from the big players - OpenAI's Operator, Anthopic's Claude Computer Use (in beta), SAP's AI Foundation on BTP and Google's Project Mariner to name but a few - but it's still difficult and expensive to deploy them usefully. Indeed, no less an authority than Google's Demis Hassabis recently cautioned that deploying

AUTHOR SPOTLIGHT
John Leonard
Research Editor
View profile

More from John Leonard
Where are all the AI agents?
NCSC calls on private researchers to help bolster critical infrastructure

Most read
01 Lovelace Report reveals £2-3.5 billion annual loss from women leaving UK tech
02 UK secretly relocates nearly 7,000 Afghans after

real-world agents is complex. “If your world model has just a 1% error rate ... by the time you’ve done those 50 or 100 steps, you’re in potentially a random place,” he said.

Blocked pipes

As Hassibis implies, the old garbage-in-garbage-out problem is compounded in an agentic setting, making data quality even more important. After all, AI agents are only as good as the data they are fed, and with many acting in real-time they are highly dependent the speed at which data architectures can feed them too. Unfortunately, most organisations of any size are characterised by data of questionable quality sitting in silos of variable accessibility connected by pipelines of dubious capacity.

Value = access

But for agents to generate real value - i.e. perform a task many times more efficiently or effectively, or to do things that were previously impossible – they need wide ranging access to internal data sources.

In a recent survey by Confluent, 84% of IT leaders agreed that AI systems must use their enterprise data to realise agentic AI’s true potential. Toran Bruce Richards, founder of UK AI firm AutoGPT, would no doubt agree.

“Here's what nobody tells you about building AI agents: the moment they actually start working is the moment they need access to everything,” he writes in OpenUK’s latest research report *From Agentic to Public Good in 2025*. “The pattern is clear: constrain access, constrain value.”

Access = security nightmare

But agents wandering about and helping themselves to whatever takes their fancy is the stuff of security nightmares.

Clearly a robust zero-trust-type architecture combined with advanced identity and access management is going to be an essential requirement if agents are going to be able to add value without leaking secrets. But this is a practical challenge in view of legacy infrastructure that wasn’t designed with zero trust in mind, the plethora of devices across which agents will be roaming, and new specialised authentication requirements.

It will require “a fundamental shift in how we think about identity, trust and security in an age where autonomous agents will become integral to our digital infrastructure," said Andrew Martin, co-founder and CEO of security company ControlPlane.

In short, there is a tension between access (value) and security.

Data oversharing

Then there’s the question of where sensitive data might end up. In one plausible scenario, given the hunger for AI training data, sensitive information may find itself in the hands of the agentic platform provider only to be regurgitated by an LLM in recognisable form somewhere down the line. Or an overzealous agent might decide that sharing data with a competitor is the fastest way to achieve its objective. The lack of transparency from AI companies does little to calm these fears - although ironically OpenUK’s report finds that businesses are slower to adopt “AI that’s open” than proprietary technology. This might be down to brand recognition, or understandable confusion about what “open AI” entails. It may change as companies wonder how much they can trust a black box.

Ethical dilemmas

As agents’ power increases, how can we be sure their aims will remain aligned with human values and prevailing norms? Who should be held responsible if they go berserk and cause real harm? The law of unintended consequences applies more and more with increasing agency, which in the corporate world manifests as more risk and compliance issues.

Will AI comply?

Shadow AI is already proving quite a handful for many cybersecurity professionals, and that’s mostly just employees trying shortcuts out of curiosity. Shoddily implemented agents acting according to skewed objectives could 10X this issue at time when regulations are evolving and guidance is thin on the ground. Are the potential rewards worth the risk? For many, the answer will be no, at least for now.

Undercover agents

All this means that agents are still operating very much undercover. But it would be wrong to conclude that nothing is going on. The prize for successful implementation could be huge.

massive MoD data breach

03 Exclusive: Google defends UK deal amid data sovereignty concerns

04 McDonald’s AI hiring bot exposed 64 million applicants’ details

05 Four arrested over M&S, Co-op and Harrods cyberattacks

"Right now we are at an early stage in seeing how agents have an impact on workflow, return on investment etc., and we don't have anything like enough data," said OpenUK CEO Amanda Brock when asked about production use cases uncovered in the research.

"Very few companies are really building agents, but the process has started."

Indeed, OpenAI is working with AutoGPT – which runs the most active open source agentic AI project on GitHub - on its own agentic experiments. The non-profit is adamant that if agentic AI is to take off, openness will be a pre-requisite over the long term for reasons of safety, security and privacy.

"The UK is uniquely positioned to lead on AI openness and public-interest technology - but only if it can address systemic challenges and harness its collaborative ecosystem," said research director Dr Jennifer Barth.

SHARE

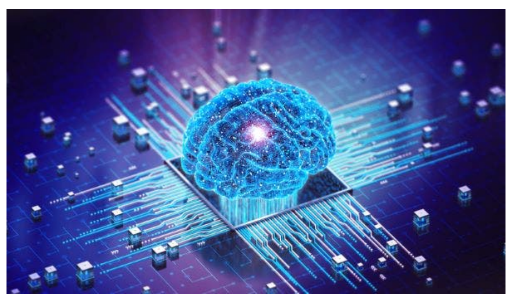
Related Topics

Artificial Intelligence | AI Ethics | Compliance | data governance | data quality | GenAI | openuk

PREVIOUS ARTICLE
EU unveils prototype age verification app and platform guidelines

NEXT ARTICLE
UK secretly relocates nearly 7,000 Afghans after massive MoD data breach

You may also like



AI

UK workers must embrace AI or risk falling behind, warns Technology Secretary 🔒

UK workers must shift from "trepidation" to "exhilaration" when it comes to AI or risk being left behind by those already engaging with the technology, ...

16 June 2025 • 4 min read

AI

Capita to roll out AI-powered recruiting this summer 🔒

Capita is the latest large employer to announce its use of agentic AI to slash recruitment timescales and budgets.

11 June 2025 • 3 min read

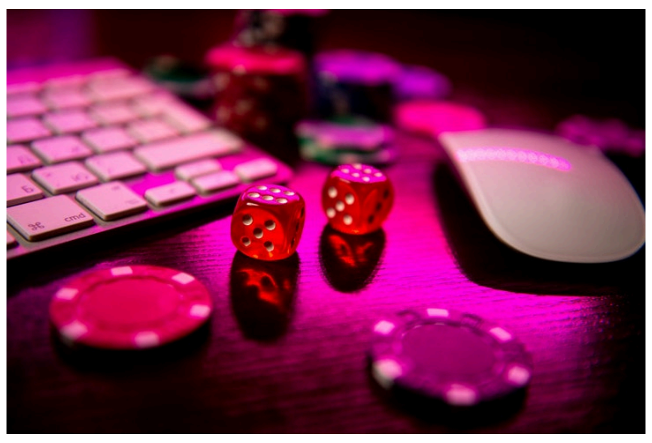
ARTIFICIAL INTELLIGENCE

UK judge warns of justice risks as lawyers cite fake AI-generated cases in court 🔒

The misuse of generative AI in legal submissions is a growing problem, and LLMs making up cases could undermine the foundations of the justice system, ...

9 June 2025 • 3 min read

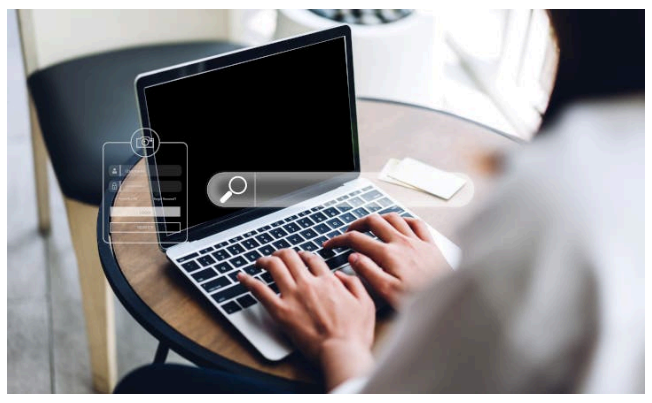
More on Artificial Intelligence



ARTIFICIAL INTELLIGENCE

Code modernisation: How bet365 cracked a gen AI conundrum 🔒

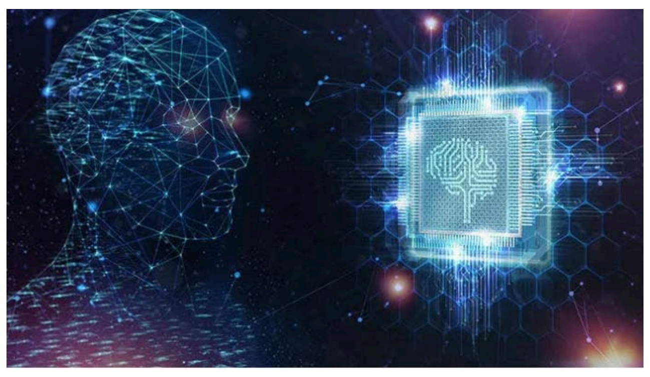
Bet365 had a breakthrough by combining metadata tagging and GraphRAG to give gen AI the all-important context it needed.



ARTIFICIAL INTELLIGENCE

OpenAI, Perplexity browsers present direct challenge to Google Chrome's dominance 🔒

OpenAI is preparing to launch a new AI-powered web browser that could significantly shake up the digital landscape dominated by Alphabet's Google Chrome.



ARTIFICIAL INTELLIGENCE

Here's how workers are using AI 🔒

As IT leaders try to get a handle on "shadow AI" - AI usage that flies under the radar of IT oversight - a new ...

THE CHANNEL CO.
computing
In partnership with MES
Hear from IT leaders tackling today's biggest challenges
SUBSCRIBE TO OUR YOUTUBE CHANNEL