LUMO INSIGHTS

Tuesday, 5 May 2026

AI agents gain power as leaders confront New risks

Executive Summary

Autonomous AI “agents” showed both impressive progress and serious warnings in the past 48 hours. Major enterprise platforms are embracing agentic automation to boost productivity, while government security agencies are urging businesses to tap the brakes and strengthen oversight. The dual message to executives: there’s huge opportunity in these technologies - but navigating their risks requires a disciplined strategy.

Security & governance: allies sound alarm

A coalition of Western cybersecurity agencies took the unusual step of issuing a joint warning about autonomous AI agents this week. Intelligence and security organisations from the US, UK, Canada, Australia, and New Zealand - collectively known as the Five Eyes - urged companies to **prioritise resilience over productivity** when deploying “agentic” AI (www.theregister.com [1]). In a coordinated advisory, they cautioned that these self-directed AI systems can behave unpredictably and dramatically **widen an organisation’s attack surface (www.theregister.com [2]).** Their message was clear: slow down and implement strong controls before letting AI systems act on their own authority.

These agencies rarely publish joint guidance on emerging technologies, which underscores their concern. They highlighted how agentic AI, by design, will plug into many tools and data sources, meaning a small glitch or malicious instruction could cascade into a much larger incident (www.theregister.com [3]). A trivial example might be an AI agent mistakenly deleting critical data or ordering erroneous transactions - mistakes at machine speed. The warning also noted that until security standards mature, organisations should **assume AI agents may “behave unexpectedly” (www.theregister.com [4])** and prepare accordingly.

What does “slow and careful adoption” mean in practice? In short: start with low-risk tasks, implement **zero-trust architecture** for AI actions, and keep human supervisors in the loop for now (www.theregister.com [5]) (www.theregister.com [6]). Notably, the agencies recommended using a “security controls checklist” for any AI agent deployment, much as one would for onboarding a privileged human employee. This includes strict permissioning (don’t let an AI have broader access than it truly needs), robust testing in sandboxed environments, and continuous monitoring of agent behaviour.

The call for governance is hitting as evidence mounts of a gap between AI agent ambitions and controls. One analysis found that roughly **60% of enterprise AI agent deployments have high-risk capabilities (like the power to execute code)** enabled, while simple two-person approval processes often fail to prevent serious mistakes. Separately, a Forrester survey found **71% of companies using AI agents lack a formal governance framework** for them - even though 64% of those same firms

plan to increase agent autonomy in the next year. In other words, most organisations are racing ahead without guardrails. For senior leaders, the mandate is to close this governance gap before scaling up agents. The Five Eyes alert should be seen as a sign that no company is exempt from these risks - and that regulators and insurers are likely watching how responsibly businesses manage this transition.

References:
[1] www.theregister.com — https://www.theregister.com/2026/05/04/five_eyes_agentic_ai_recommendations/#:~:text=Eyes%20spook%20shops%20warn%20rapid,delivered%20that%20position%20last%20Friday
[2] www.theregister.com — https://www.theregister.com/2026/05/04/five_eyes_agentic_ai_recommendations/#:~:text=AI,additional%20avenues%20of%20exploitation%2C%E2%80%9D%20the
[3] www.theregister.com — https://www.theregister.com/2026/05/04/five_eyes_agentic_ai_recommendations/#:~:text=unexpectedly%20The%20thrust%20of%20the,crafts%20a%20seemingly%20innocuous%20prompt
[4] www.theregister.com — https://www.theregister.com/2026/05/04/five_eyes_agentic_ai_recommendations/#:~:text=AI,agentic%20AI%20system%20widens%20the
[5] www.theregister.com — https://www.theregister.com/2026/05/04/five_eyes_agentic_ai_recommendations/#:~:text=Canada%20Simon%20Sharwood%20Mon%204,delivered%20that%20position%20last%20Friday
[6] www.theregister.com — https://www.theregister.com/2026/05/04/five_eyes_agentic_ai_recommendations/#:~:text=AI,additional%20avenues%20of%20exploitation%2C%E2%80%9D%20the

Enterprise platforms: agent-first moves

Despite the cautionary tone from regulators, technology heavyweights are rapidly expanding the capabilities of AI agents in enterprise environments. In the last two days, several major platform providers rolled out new “agent-first” features - essentially inviting AI to take the wheel inside their systems.

Salesforce, for example, unveiled an initiative called **Headless 360** to make its entire Customer 360 platform accessible to AI agents via APIs, command-line interfaces, and approved toolkits, with no graphical user needed (venturebeat.com [1]). In practice, this means a capable AI could perform any operation a human user can in Salesforce - from updating CRM records to generating reports - all programmatically. Industry analysts note that Salesforce is positioning itself to become a “system of execution” for AI-driven workflows, not just a system of record (www.cio.com [2]). By exposing over 100 new functions for direct machine use and baking in safety features (like requiring confirmation for risky actions), Salesforce is betting that companies want their critical business processes to be as easy for AI to navigate as a website.

Not to be outdone, **Microsoft’s Agent 365** officially hit general availability this week, signalling that Redmond believes autonomous agents are ready for large-scale enterprise use. Agent 365 is a unified control plane for managing and monitoring AI agents across Microsoft’s ecosystem - from Office 365 and Teams to Azure cloud services (msftnewsnow.com [3]) (msftnewsnow.com [4]). It integrates with familiar tools like Microsoft Entra ID (for identity and access control), Defender security, and Intune device management to track everything AI agents do as first-class “digital employees.” By including Agent 365 in its new top-tier Office E7 license and integrating it with competing clouds like AWS and Google (msftnewsnow.com [5]) (msftnewsnow.com [6]), Microsoft is acknowledging both the demand for cross-platform agent orchestration and the need to control the "shadow AI" that might otherwise spread unchecked. This is a direct play to help CIOs deploy autonomous workflows with confidence that security and compliance won’t be left behind.

Even core infrastructure players are embracing agent autonomy. Cloudflare, a major cloud networking firm, just expanded its platform to let AI systems deploy and manage applications on its global edge network automatically - a task that used to require human DevOps teams. Cloudflare’s aim is to support “millions of autonomous, long-running agents” on its network, moving beyond simple chatbots to AI that actually runs software at scale (www.cloudflare.com [7]) (www.cloudflare.com [8]). And Google, for its part, has been championing open agent interoperability: its new **Agent-to-Agent

(A2A)** communication protocol (now in production use at 150 companies) allows different vendors' AI agents to coordinate tasks with each other seamlessly (thenextweb.com [9]). That interoperability effort is being governed under the Linux Foundation's **Agentic AI Foundation**, which brings together competitors like Google, OpenAI, Microsoft, and Anthropic to set common standards (thenextweb.com [10]). Taken together, these moves show how fast the ecosystem is moving: enterprises will soon have agent capabilities integrated into everything from CRM systems to IT infrastructure - and even the means for those agents to talk to each other across platforms.

References:
[1] venturebeat.com — https://venturebeat.com/technology/salesforce-launches-headless-360-to-turn-its-entire-platform-into-infrastructure-for-ai-agents#:~:text=Headless%20360%20,more%20than%20100%20new%20tools
[2] www.cio.com — https://www.cio.com/article/4159536/salesforce-launches-headless-360-to-support-agent-first-enterprise-workflows.html#:~:text=layer%20for%20enterprise%20AI%20agents,added%2C%20Salesforce%20is%20trying%20to
[3] msftnewsnow.com — https://msftnewsnow.com/microsoft-agent-365-ga-control-plane-ai-agents/#:~:text=general%20availability%20of%20Microsoft%20Agent,multiply%20across%20their%20environment%20without
[4] msftnewsnow.com — https://msftnewsnow.com/microsoft-agent-365-ga-control-plane-ai-agents/#:~:text=What%20Microsoft%20Agent%20365%20Is
[5] msftnewsnow.com — https://msftnewsnow.com/microsoft-agent-365-ga-control-plane-ai-agents/#:~:text=Agent%20365%20is%20now%20generally,month%2C%20targeting%20the%20people%20who
[6] msftnewsnow.com — https://msftnewsnow.com/microsoft-agent-365-ga-control-plane-ai-agents/#:~:text=match%20at%20L143%20Microsoft%20is,guessing%20about%20what%E2%80%99s%20running%20where
[7] www.cloudflare.com — https://www.cloudflare.com/press/press-releases/2026/cloudflare-expands-its-agent-cloud-to-power-the-next-generation-of-agents/#:~:text=cloud%20company%2C%20is%20today%20expanding,generation%20of%20AI%20focused%20on
[8] www.cloudflare.com — https://www.cloudflare.com/press/press-releases/2026/cloudflare-expands-its-agent-cloud-to-power-the-next-generation-of-agents/#:~:text=infrastructure%2C%20compute%2C%20deployment%20and%20security,founder%20and%20CEO%20of
[9] thenextweb.com — https://thenextweb.com/news/google-cloud-next-ai-agents-agentic-era#:~:text=technology%20partners%2C%20has%20reached%20150,using%20cryptographic%20signatures%20for%20domain
[10] thenextweb.com — https://thenextweb.com/news/google-cloud-next-ai-agents-agentic-era#:~:text=technology%20partners%2C%20has%20reached%20150,using%20cryptographic%20signatures%20for%20domain

Technologies enabling autonomy

Why is all this happening now? One driver is that the AI models themselves are getting dramatically more powerful and efficient. OpenAI's new **GPT-5.5** and Anthropic's **Claude Opus 4.7** - both released in the past week - take another leap in reasoning and code-writing ability. These models can juggle longer instructions and perform more complex operations than their predecessors. Critically, they are being optimised to work across multiple tools and systems, which is the key for true "autonomous" workflows. For instance, Adobe just introduced its **Firefly AI** assistant that spans the entire Creative Cloud suite. You can now simply tell Adobe's AI what you need - say, create a marketing video - and it will generate assets in Photoshop, import them into Premiere, edit according to your brief, and produce a finished video in minutes. That kind of multi-step, multi-app capability was the stuff of demos a year ago; today it's in the hands of early users.

Another factor is the rapid improvement in inference efficiency - the speed and cost of running these big models in real applications. It's often said that 2023 was about making AI "bigger," while 2026 is about making it **cheap and fast** enough to deploy everywhere. This week saw evidence of that shift. Specialised AI chips and model optimisation techniques are slashing the cost of running agents. One new system reportedly cut the cloud cost of an AI coding agent by 17× by swapping in a more efficient model mid-task. And industry-wide, chipmakers and cloud providers are racing to reduce the per-query cost of advanced AI. The takeaway for enterprises is that the barrier to entry - the cost of letting an AI system work continuously on real business problems - is dropping quickly.

As AI agents gain these new powers, vendors are also adding safeguards. For example, robust **deterministic guardrails** (hard-coded rules about what an agent can and cannot do) are being built into enterprise agent platforms to guarantee safety on critical tasks (www.salesforce.com [1]). And "context engineering" - carefully controlling what data an AI can access and remember - is emerging

as a new discipline to keep agents on-mission (www.salesforce.com [2]). These developments suggest that technical progress in 2026 isn't just about raw intelligence; it's also about making smart machines that leaders can actually trust with business operations.

References:
[1] www.salesforce.com — https://www.salesforce.com/blog/ai-agent-trends-2026/#:~:text=trends%20shaping%20enterprise%20AI,Early
[2] www.salesforce.com — https://www.salesforce.com/blog/ai-agent-trends-2026/#:~:text=right%20thing%20to%20agents%20that,under%20which%20the%20question%20is

Adoption, talent, and the trust gap

Amid the excitement, the on-the-ground reality is that most enterprises are still in the early days of using AI agents - and many are understandably cautious. Surveys indicate that while roughly four out of five companies have tried some form of AI agent, only a tiny fraction (just 2%) have fully rolled them out across the organisation. The rest are stuck in experimentation mode. The reason most often cited by executives? Lack of confidence. In one recent poll, **55% of business leaders said they worry AI agents aren't reliable enough,** prone to errors or unpredictable actions. This credibility gap is a major hurdle: even if the technology is available, it won't transform operations unless people trust it.

Building that trust will require a concerted effort on multiple fronts. First, early wins are crucial - identifying low-hanging fruit where agents can quickly save time and prove their value without jeopardising the business. We're already seeing such wins: companies that have embraced AI agents are processing significantly more work with the same staff, reporting productivity boosts above 50% in some cases. A case study from the insurance industry showed an 80% reduction in tedious form-filling after deploying an underwriting assistant, freeing human brokers to spend more time with clients. Success stories like that will help convince sceptical stakeholders that these tools can deliver real benefits.

Second, organisations are recognising that they need new skills and roles to safely harness AI agents. A few years ago, job titles like **"AI Ops Manager" or "Agent Supervisor"** would have drawn blank stares; now, forward-thinking companies are hiring for these positions (www.salesforce.com [1]). These people oversee the behaviour and performance of AI systems, much like a manager would supervise human team members. A formal **Agent Development Lifecycle** is emerging, from design and testing to ongoing monitoring and "AI QA" of outputs (www.salesforce.com [2]). Leaders should ensure they have (or are developing) this kind of internal capability - whether by upskilling existing staff with AI training programmes or bringing in new talent.

Finally, it's increasingly apparent that delaying too long carries its own risk. Competitors who move quickly are not only enjoying efficiency gains, but may also start capturing market share by reallocating human effort to higher-value work. One industry briefing this week put it bluntly: the advantage goes to **whoever deploys effective agents first.** The message for executives is to neither rush in blindly nor cling to total caution. The winning approach will be to **embrace these new agent tools with a clear plan**: get your team educated (even Google is offering a free crash course on building AI agents), put strong policies and oversight in place, and start integrating agents into operations in a controlled way. Those that do so are likely to see outsized gains - and those that don't could be left doing manual work while others forge ahead.

References:
[1] www.salesforce.com — https://www.salesforce.com/blog/ai-agent-trends-2026/#:~:text=maps%20the%20full%20lifecycle%20from,more%20AI%20news%20and%20insights
[2] www.salesforce.com — https://www.salesforce.com/blog/ai-agent-trends-2026/#:~:text=happens%20after%20deployment,a%20core%20part%20of%20enterprise

KEY TAKEAWAY

Move fast, but don’t break things. The potential productivity upsides of AI agents are too significant to ignore - but the warning from global security experts is clear. Senior leaders should champion agile adoption of autonomous AI while enforcing strong governance, oversight, and upskilling to manage the risks.

Sources

Five Eyes warn agentic AI is too dangerous for rapid rollout
https://www.theregister.com/2026/05/04/five_eyes_agentic_ai_recommendations/

Salesforce launches Headless 360 to support agent-first enterprise workflows
https://www.cio.com/article/4159536/salesforce-launches-headless-360-to-support-agent-first-enterprise-workflows.html

Microsoft Agent 365 Goes Live as Company Unveils E7 Suite
https://redmondmag.com/articles/2026/05/01/microsoft-agent-365-goes-live-as-company-unveils-e7-suite.aspx

8 Ways AI Agents Are Evolving in 2026
https://www.salesforce.com/blog/ai-agent-trends-2026/