LUMO INSIGHTS

Tuesday, 19 May 2026

AI agents take charge: New tech and guardrails for enterprises

Executive Summary

In the last two days, AI agents and autonomous workflows have leapt forward in capability and oversight. From a first-ever CPU built specifically for AI agents to official warnings about new security gaps, the enterprise landscape is shifting from experimental AI towards scaled, governed automation.

Specialised hardware for autonomous AI

This week marked a milestone in AI infrastructure: NVIDIA began shipping its new Vera processor - the first CPU designed specifically for AI agents - to leading AI labs. The 88-core Vera processors, hand-delivered to organisations like OpenAI, Anthropic and even SpaceX's AI division, are purpose-built to handle the orchestration and tool-calling demands of autonomous agents at high speed. In contrast to traditional chips that merely serve up AI model predictions, these CPUs are optimised for the complex, non-linear workloads that emerge when AI models start to take actions in the real world (blogs.nvidia.com [1]).

NVIDIA's move underscores how seriously the industry is taking agentic AI. CEO Jensen Huang has compared Vera's launch to a new 'CPU moment' for the era of AI, as companies push beyond simple question-answering bots to AI "co-workers" that can plan, execute, and adapt dynamically. The Vera architecture focuses on exactly those tasks: it can juggle long-running conversations, call external applications in real time, and manage massive context windows without bogging down. Early reports claim that using Vera for AI agent workloads can cut processing costs by up to 90% while significantly boosting speed.

For enterprises, the takeaway is that the computing foundations for autonomous workflows are rapidly evolving. Specialised hardware like Vera will likely appear in cloud data centres soon, meaning the next generation of AI platforms will run faster and cheaper. Leaders should keep an eye on how their technology vendors incorporate this kind of innovation - it could unlock more ambitious AI projects (and cost savings) sooner than expected.

References:

[1] blogs.nvidia.com — https://blogs.nvidia.com/blog/vera-cpu-delivery/#:~:text=NVIDIA%E2%80%99s%20next%20multi,delivered%20them.%20%E2%80%9CAgentic%20AI%20is

Owning the agent ecosystem

A quiet battle is unfolding over who will control the building blocks of agentic AI. Yesterday, Anthropic - one of the leading AI model providers - made a strategic move that sent a clear message. It bought a startup called Stainless, which provided a clever way for developers to generate and update software development kits (SDKs) for AI models automatically (techcrunch.com [1]). Engineers at OpenAI,

Google, and even Cloudflare had been using Stainless to simplify their AI integrations (techcrunch.com [2]). By snapping it up (for a reported $300 million (techcrunch.com [3])) and planning to cut off its services to competitors (techcrunch.com [4]), Anthropic is effectively pulling up the ladder behind itself.

This acquisition highlights an emerging trend: major AI players are racing to build end-to-end ecosystems for enterprise AI agents. They're not just offering powerful models; they want to own the whole stack - from custom silicon to developer tools to managed platforms. OpenAI, for instance, has been rolling out its own suites for enterprise AI agents, aiming to bind customers to its ecosystem. Now Anthropic is countering by locking down a piece of the "plumbing" that every AI application relies on. For enterprise leaders, this intensifying competition could lead to a more fragmented landscape, where interoperability becomes a challenge and vendor lock-in risks grow.

At the same time, new solutions are appearing to keep the ecosystem flexible. On one front, startups like WaveSpeed are tackling the complexity head-on by offering unified interfaces to many AI models: instead of getting tied to a single vendor, companies can use one API to tap into a whole library of models from different providers and even different modalities (text, image, audio) (pr.capitalpress.com [5]) (www.prweb.com [6]). And beyond the commercial offerings, the open-source community is busy developing standards to ensure AI agents remain portable and extensible - for example, an emerging secure registry of vetted agent 'skills' for coding assistants aims to prevent unsafe plug-ins. The upshot: CIOs and tech strategists should decide whether to embrace a single vendor's all-in-one agent ecosystem or adopt a more open, multi-model approach. The former can offer tight integration but at the cost of dependency, while the latter preserves flexibility and leverage at the expense of some extra complexity.

References:

[1] techcrunch.com — https://techcrunch.com/2026/05/18/anthropic-has-acquired-the-dev-tools-startup-used-by-openai-google-and-cloudflare/#:~:text=development%20kits%2C%20or%20SDKs%20%E2%80%94,popular%20tool%20because%20the%20platform

[2] techcrunch.com — https://techcrunch.com/2026/05/18/anthropic-has-acquired-the-dev-tools-startup-used-by-openai-google-and-cloudflare/#:~:text=Anthropic%20has%20acquired%20the%20dev,May%2018%2C%202026%20Anthropic%20announced

[3] techcrunch.com — https://techcrunch.com/2026/05/18/anthropic-has-acquired-the-dev-tools-startup-used-by-openai-google-and-cloudflare/#:~:text=and%20Google,a%20key%20infrastructure%20supplier%20out

[4] techcrunch.com — https://techcrunch.com/2026/05/18/anthropic-has-acquired-the-dev-tools-startup-used-by-openai-google-and-cloudflare/#:~:text=than%20%24300%20million,and%20extend%20them%20however%20they

[5] pr.capitalpress.com — https://pr.capitalpress.com/article/WaveSpeed-Expands-Unified-LLM-API-With-Access-to-GPT-Claude-Gemini-and-260-Models?storyId=6a09cab9d6b3ec3c381af81c#:~:text=language%20models%20with%20a%20broader,AI

[6] www.prweb.com — https://www.prweb.com/releases/wavespeed-expands-unified-llm-api-with-access-to-gpt-claude-gemini-and-260-models-302773456.html#:~:text=unified%20AI%20model%20platform%20that,WaveSpeed%20is%20used%20by

Agents in action: real-world workflow automation

Perhaps the most convincing evidence that AI agents have matured is the real-world impact they're already delivering. Adoption is now widespread - one survey found 96% of enterprises are using AI agents in some capacity (www.prnewswire.com [1]) - and about 80% of organisations say these projects are yielding measurable returns on investment (claude.com [2]). In other words, autonomous agents are no longer just science experiments; they are starting to pay off in practice. The improvements can be dramatic: a major manufacturer, for instance, reduced a critical production optimisation from six weeks to one day by letting an AI agent handle it (openai.com [3]). And a large energy company reported a 5% increase in output (worth over $1 billion annually) thanks to agent-driven adjustments in its operations (openai.com [4]).

In healthcare, AI agents are tackling stubborn operational bottlenecks. One hospital recently introduced a "virtual dispatch" agent to fill last-minute nursing shift vacancies via automated SMS outreach. The result: the agent finds a replacement in under five minutes - a task that previously took managers up to five hours of calls and messages (hitconsultant.net [5]). That speed translated into a

25% reduction in overtime costs at the hospital, relieving staffing pressure and burnout at a time of chronic nurse shortages (hitconsultant.net [6]). It's a clear example of how agent automation can solve problems that humans used to just work around.

Even knowledge and creative work is being transformed. In marketing, for example, Skai's newly launched platform lets swarms of AI agents run digital ad campaigns with minimal human intervention (skai.io [7]). These agent squads continuously monitor campaign metrics across channels, diagnosing issues and tweaking budgets in real time to hit performance targets (skai.io [8]). Human marketers still set the strategy and creative direction, but much of the tedious day-to-day optimisation - tasks that once consumed countless hours - can now happen autonomously at machine speed. The early results are promising: faster campaign adjustments, more agility in responding to market changes, and marketers freed to focus on strategy and creative ideas (skai.io [9]). Similar experiments are underway in customer service, finance, and other fields as companies test how far they can safely hand off routine workflows to AI.

References:

[1] www.prnewswire.com — https://www.prnewswire.com/apac/news-releases/agentic-ai-goes-mainstream-in-the-enterprise-but-94-raise-concern-about-sprawl-outsystems-research-finds-302739251.html#:~:text=experimentation%20to%20execution,critical%20operations.%20This%20shift

[2] claude.com — https://claude.com/blog/how-enterprises-are-building-ai-agents-in-2026#:~:text=next%20year,Thomson%20Reuters%20uses%20Claude%20to

[3] openai.com — https://openai.com/index/introducing-openai-frontier/#:~:text=production%20optimization%20work%20from%20six,which%20adds%20over%20a%20billion

[4] openai.com — https://openai.com/index/introducing-openai-frontier/#:~:text=production%20optimization%20work%20from%20six,which%20adds%20over%20a%20billion

[5] hitconsultant.net — https://hitconsultant.net/2026/05/14/chromie-health-raises-2m-ai-nurse-dispatch/#:~:text=Ventures,chaos%20through%20phone%20calls%2C%20group

[6] hitconsultant.net — https://hitconsultant.net/2026/05/14/chromie-health-raises-2m-ai-nurse-dispatch/#:~:text=Ventures,chaos%20through%20phone%20calls%2C%20group

[7] skai.io — https://skai.io/press-releases/skai-unveils-the-industrys-first-agent-native-operating-system-for-marketing/#:~:text=Skai%20Unveils%20the%20Industry%E2%80%99s%20First,native

[8] skai.io — https://skai.io/press-releases/skai-unveils-the-industrys-first-agent-native-operating-system-for-marketing/#:~:text=environment,agent%20operates%20within%20guardrails%20defined

[9] skai.io — https://skai.io/press-releases/skai-unveils-the-industrys-first-agent-native-operating-system-for-marketing/#:~:text=into%20a%20single%20operating%20environment,At%20the%20center

Governing the Age of AI Co-Workers

As businesses deploy more autonomous agents, a new challenge is coming into focus: how to govern these AI co-workers so they remain reliable and safe. This week, the U.S. National Institute of Standards and Technology (NIST) released a report synthesising industry feedback on AI agent security (www.nist.gov [1]). The consensus is sobering. Nearly all respondents agreed that AI agents introduce unique security threats, from the potential for an agent's broad permissions to be misused by malicious insiders to the possibility of automated systems making mistakes at a scale and speed humans never could. NIST's analysis concludes that while traditional cybersecurity principles (access controls, identity management, monitoring, etc.) remain crucial, they will require adaptation to address this new class of risk (www.nist.gov [2]).

The governance gap is becoming impossible to ignore. Adoption of agentic AI is far outpacing the guardrails: according to one survey, 94% of organisations are worried that AI "sprawl" is increasing faster than they can manage it, adding complexity, technical debt and risk to the business (www.prnewswire.com [3]). We've already seen hints of the dangers. In a joint advisory, intelligence agencies from the US, UK, Canada, Australia and New Zealand (the "Five Eyes" alliance) recently warned that rushing out agentic AI without proper controls will likely result in serious incidents. They gave examples like an AI agent asked to install software patches that, due to an overly broad mandate, also wipes critical security logs, blinding IT to an intrusion (www.theregister.com [4]). Without strict policies and oversight, autonomous systems can amplify vulnerabilities or create new

ones.

For senior leaders, the mandate is clear: treat AI agents as both a strategic asset and a potential liability. Just as companies enforce controls around human employees or outsourcing partners, they need equally robust frameworks for AI agents. That means maintaining an inventory of every deployed agent and its role, limiting each agent's access permissions to the minimum necessary, monitoring agent decisions and actions in real time, and having emergency off switches and incident response plans in place for when things go wrong. Regulators and standards bodies are beginning to weigh in - and early guidance is likely to mirror these best practices. Ensuring strong governance isn't a brake on innovation; it's what will enable organisations to scale up the use of autonomous workflows with confidence.

References:

[1] www.nist.gov — https://www.nist.gov/publications/summary-analysis-responses-request-information-regarding-security-considerations-ai#:~:text=government%20could%20offer%20to%20facilitate,for%20government%20action%2C%20including%20offering
[2] www.nist.gov — https://www.nist.gov/publications/summary-analysis-responses-request-information-regarding-security-considerations-ai#:~:text=government%20could%20offer%20to%20facilitate,Commenters
[3] www.prnewswire.com — https://www.prnewswire.com/apac/news-releases/agentic-ai-goes-mainstream-in-the-enterprise-but-94-raise-concern-about-sprawl-outsystems-research-finds-302739251.html#:~:text=adoption%20accelerates%2C%20governance%20is%20struggling,only%20a%20small%20fraction%20of
[4] www.theregister.com — https://www.theregister.com/security/2026/05/04/five-eyes-warn-agentic-ai-is-too-dangerous-for-rapid-rollout/5229103#:~:text=document%20offers%20the%20example%20of,logs%20because%20its%20permissions%20allow

Key Statistics

- 96% of surveyed enterprises are already using AI agents in some capacity ([www.prnewswire.com](https://www.prnewswire.com/apac/news-releases/agentic-ai-goes-mainstream-in-the-enterprise-but-94-raise-concern-about-sprawl-outsystems-research-finds-302739251.html#:~:text=experimentation%20to%20execution,critical%20operations.%20This%20shift))
- 80% of organisations report their AI agent initiatives are delivering measurable ROI ([claude.com](https://claude.com/blog/how-enterprises-are-building-ai-agents-in-2026#:~:text=next%20year,Thomson%20Reuters%20uses%20Claude%20to))
- NVIDIA's Vera CPU is an 88-core chip with 1.2 /TB/s memory bandwidth, purpose-built for AI agent workloads ([wccftech.com](https://wccftech.com/nvidia-hand-delivers-first-vera-cpus-to-anthropic-openai-spacex-oracle-agentic-ai-era/#:~:text=agent%20sandboxing%2C%20long,include%20OpenAI%2C%20SpaceX%2C%20Anthropic%2C%20and))
- An AI scheduling agent filled hospital shift vacancies in under 5 minutes (versus ~5 hours manually) ([hitconsultant.net](https://hitconsultant.net/2026/05/14/chromie-health-raises-2m-ai-nurse-dispatch/#:~:text=Ventures,chaos%20through%20phone%20calls%2C%20group))

KEY TAKEAWAY

Autonomous agents are rapidly moving from experiments to core operations. Leaders must move beyond pilots, ready their infrastructure and governance, and integrate agent-led workflows now - or risk falling behind.

Sources

Summary Analysis of Responses to the Request for Information Regarding Security Considerations for AI Agents
https://www.nist.gov/publications/summary-analysis-responses-request-information-regarding-security-considerations-ai
Vera Arrives: NVIDIA's First CPU Built for Agents Lands at Top AI Labs
https://blogs.nvidia.com/blog/vera-cpu-delivery/

Anthropic has acquired the dev tools startup used by OpenAI, Google, and Cloudflare
https://techcrunch.com/2026/05/18/anthropic-has-acquired-the-dev-tools-startup-used-by-openai-google-and-cloudflare/

Agentic AI Goes Mainstream in the Enterprise, but 94% Raise Concern About Sprawl (OutSystems Research)
https://www.prnewswire.com/apac/news-releases/agentic-ai-goes-mainstream-in-the-enterprise-but-94-raise-concern-about-sprawl-outsystems-research-finds-302739251.html

How enterprises are building AI agents in 2026 | Claude Blog
https://claude.com/blog/how-enterprises-are-building-ai-agents-in-2026

WaveSpeed Expands Unified LLM API With Access to GPT, Claude, Gemini and 260+ Models
https://www.prweb.com/releases/wavespeed-expands-unified-llm-api-with-access-to-gpt-claude-gemini-and-260-models-302773456.html

Skai Unveils the Industry's First Agent-Native Operating System for Marketing
https://skai.io/press-releases/skai-unveils-the-industrys-first-agent-native-operating-system-for-marketing/

Chromie Health Raises $2M to Automate Hospital Operations via AI Agent 'Dispatch'
https://hitconsultant.net/2026/05/14/chromie-health-raises-2m-ai-nurse-dispatch/

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