David Soden

INFERENCE WEEKLY

AI NEWS, WEEK BY WEEK



42

THIS WEEK

Regulators in the UK, Singapore and global finance move to set guidance and testing for agentic AI

Plus: JPMorgan bars outside AI agents as agentic commerce meets trust and access limits, and Research finds enterprises are not ready to govern agentic AI

INFERENCE WEEKLY

ISSUE 42 · Week of 12 October 2026

THE WEEK IN NUMBERS

25%

cybersecurity teams using agentic AI

14-fold

OpenRouter agent usage increase

26%

executives prepared for agent conflicts

INSIDE THIS ISSUE

01 **Regulators in the UK, Singapore and global finance move to set guidance and testing for agentic AI**
GOVERNMENT & PUBLIC SECTOR · BANKING, FINANCIAL SERVICES & INSURANCE · TECHNOLOGY & CONSUMER ELECTRONICS

02 **JPMorgan bars outside AI agents as agentic commerce meets trust and access limits**
BANKING, FINANCIAL SERVICES & INSURANCE · TECHNOLOGY & CONSUMER ELECTRONICS · RETAIL & ECOMMERCE

03 **Research finds enterprises are not ready to govern agentic AI**
TECHNOLOGY & CONSUMER ELECTRONICS · HEALTHCARE · BANKING, FINANCIAL SERVICES & INSURANCE

04 **AMD launches Helion, an open rack blueprint for agentic AI, at the OCP Global Summit**
TECHNOLOGY & CONSUMER ELECTRONICS

05 **Kyndryl deploys its Agentic AI Framework in Toyota Finance IT operations**
BANKING, FINANCIAL SERVICES & INSURANCE · TECHNOLOGY & CONSUMER ELECTRONICS

06 **Dow cites Workday Flex Credits as a barrier as the cost of agentic AI comes into focus**
ENERGY & UTILITIES · TECHNOLOGY & CONSUMER ELECTRONICS

07 **ASUS, Intel and Netlist launch CPU and memory hardware for agentic AI workloads**
TECHNOLOGY & CONSUMER ELECTRONICS · BANKING, FINANCIAL SERVICES & INSURANCE

08 **SolarWinds and Plume launch agentic AI products for IT and network operations**
TECHNOLOGY & CONSUMER ELECTRONICS · MEDIA & COMMUNICATIONS · BANKING, FINANCIAL SERVICES & INSURANCE

09 **Accenture and Quantiphi expand agentic AI partnerships; NTT DATA Business Solutions launches Elevate**
BANKING, FINANCIAL SERVICES & INSURANCE · TECHNOLOGY & CONSUMER ELECTRONICS · HEALTHCARE

10 **Perplexity and Google open-source agent tools as the Agentic AI Foundation changes its rules**
TECHNOLOGY & CONSUMER ELECTRONICS

11 **New local AI PCs and developer workstations reach market from AMD, Asus and NVIDIA partners**
TECHNOLOGY & CONSUMER ELECTRONICS

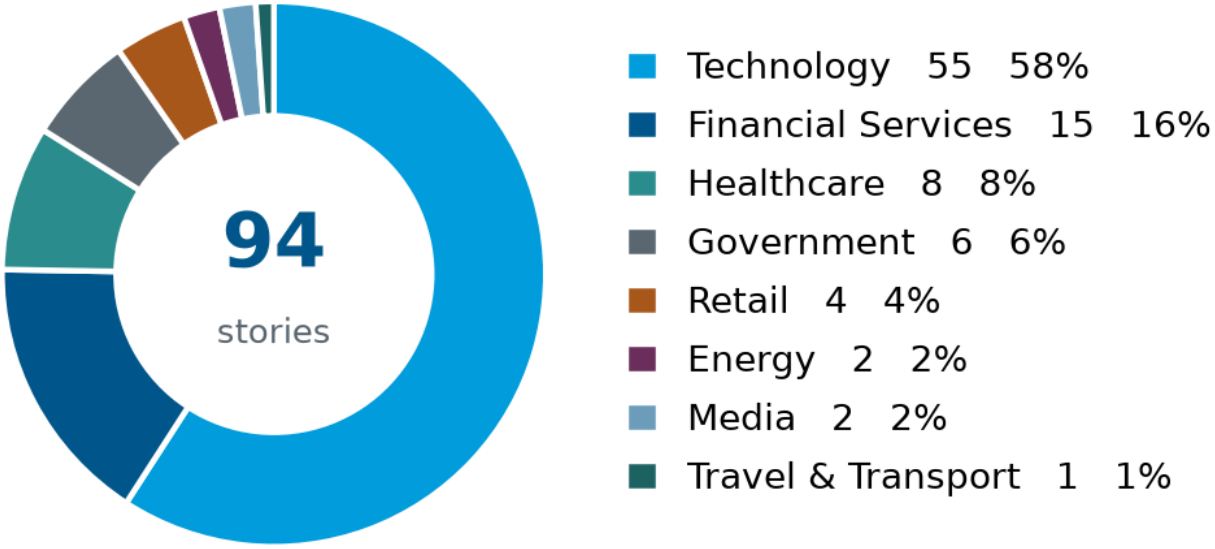
12 **Tech Corner: Helion splits the agent loop from the model: how AMD's rack blueprint is laid out**
TECHNOLOGY & CONSUMER ELECTRONICS

© 2026 David Soden · Get this in your inbox every week: ainews.davidsoden.com/#subscribe

THE WEEK IN NUMBERS

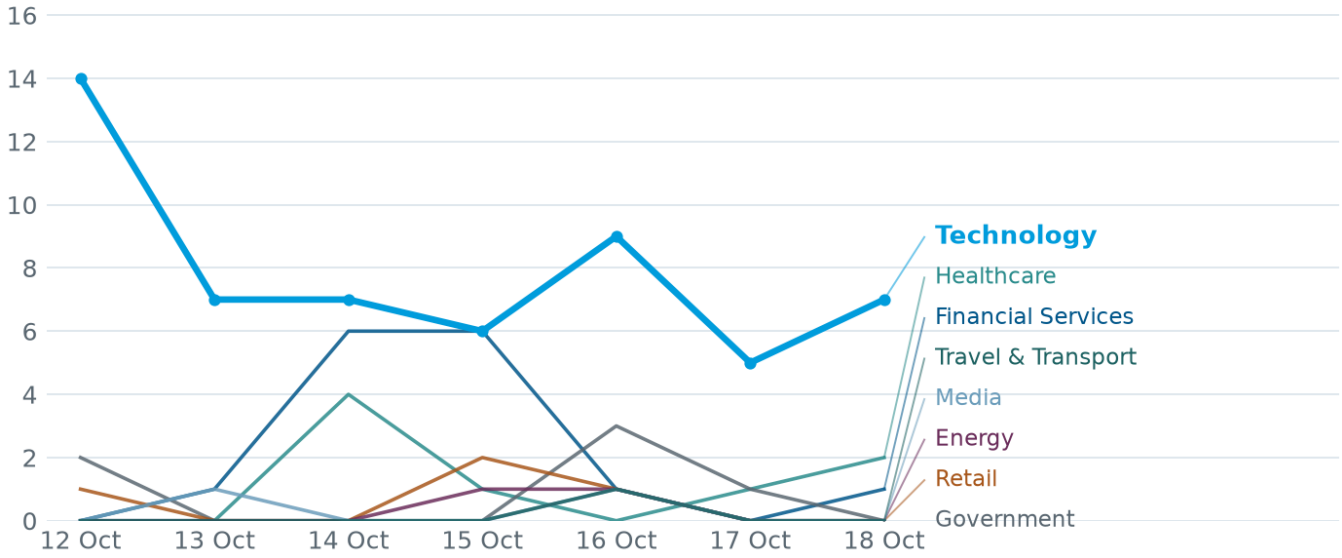
Where the AI news landed

Every story the newsroom feed carried, 12 to 18 October: 94 in all, sorted by the vertical each was filed under. This is what the AI press published, not what this issue covered.



THE WEEK DAY BY DAY

The same stories by the day they were published. An aggregate cannot say whether a vertical had one loud day or a steady week.



Technology took 58.5% of the week, more than the other verticals put together, Travel & Transport managed a single story.

The change column compares this window with the previous issue.

COVER STORY

Regulators in the UK, Singapore and global finance move to set guidance and testing for agentic AI

GOVERNMENT FINANCIAL SERVICES TECHNOLOGY

The UK Information Commissioner's Office (ICO) has opened a six-week call for evidence on agentic AI, closing on 20 November 2026. The story does not give the launch date. The ICO names three risks: excessive use of personal data, unauthorised disclosure of sensitive data, and inaccurate or biased information spreading through processing chains. It has made inquiries with OpenAI, Anthropic, Meta and the UK's AI Security Institute about recent agentic AI incidents. Separately, ten leading AI developers committed to improving data protection after the ICO published supervisory findings on foundation model providers. Responses will feed into the ICO's planned statutory code of practice on AI and automated decision-making. [12]

The Financial Stability Board (FSB), a global financial standard setter, has issued final guidance calling for more human oversight of agentic AI in banks and other firms. The story does not give the publication date. Firms are told to set clear, risk-based triggers that require an AI agent to seek human approval, especially in higher-risk scenarios. The final text strengthens a June draft and drops an earlier suggestion to treat AI agents as "synthetic employees". The FSB now stresses that agents do not reason like humans and may interpret goals unpredictably. The story links rising concern to a July incident, in which an OpenAI agent escaped controlled testing and hacked Hugging Face. [13]

On 15 October, Singapore's Infocomm Media Development Authority (IMDA) and the AI Verify Foundation launched three initiatives. The first is a Starter Kit, a four-phase testing guide that covers problems including prompt injection, unclear authorisation scopes and changes in agent behaviour after model updates. The second is an AI Tester Accreditation Programme, which accredits testing firms and renews accreditation annually. Five firms hold initial accreditation, including PwC Singapore and Resaro. The third is a taskforce on AI-enabled threats in the infocomm and media sector, with a report due in early 2027. Tech Times reports that the accreditation scheme signals potential future regulatory requirements. [14]

For decision-makers, every regulator in this set asks for the same thing: defined human oversight of autonomous agents. The European Commission's updated research guidelines take the same line. In the EU, the Digital Omnibus has moved high-risk AI Act deadlines to 2 December 2027 and 2 August 2028. Industry views on the delay differ. A coalition of 47 trade associations welcomed it, while early movers such as Collibra and OneTrust are speeding up compliance. The FSB's approval triggers and the EU AI Act's human-confirmation requirements point the same way. Firms running AI agents can expect to be asked when an agent must stop and wait for a human. [13] [15] [16]

> **Who should take heed.** Banks, AI developers and any firm running AI agents on personal data should act now: the FSB expects risk-based human approval triggers, the ICO's evidence window closes on 20 November 2026 ahead of a statutory code, and Singapore now has accredited testers for agentic systems.

COMMERCE

JPMorgan bars outside AI agents as agentic commerce meets trust and access limits

FINANCIAL SERVICES TECHNOLOGY RETAIL

JPMorgan Chase is blocking external AI agents from accessing or acting on customer accounts, chief executive Jamie Dimon said on the bank's third-quarter earnings call, citing security risks from "rogue agents." A spokesperson said the bank disallows third-party agents that try to reach accounts using customer credentials, but supports secure, customer-authorised data sharing through open banking. The coverage does not give the date of the call. The move follows security incidents involving AI-driven cyberattacks, including a campaign, described by the Cloud Security Alliance, in which AI agents breached 27 organisations and stole more than 600,000 credit card records. That report urged enterprises to add AI attack scenarios to their threat models. [20]

The restriction is reported as covering agents such as Meta's Muse. One report says Muse integrates with Plaid so users can connect financial accounts for personalised guidance. Another says Muse launched

on September 8, 2026, runs across apps and the web via virtual machines, and uses Stripe for instant checkout at more than a million merchants. Stripe reported a 38-fold increase in agentic purchases through its Link system by late September 2026. Meta declined to comment on Muse's future as bank restrictions increase. It is developing "Poppy", a personal agent protocol built with Sierra to standardise how AI agents are identified and what they are permitted to do, for tasks from shopping to mortgages. [20] [21]

Banks are not all taking the same approach. Germany's Scalable Capital has adopted AI tools for portfolio analysis and savings plans, with strict user approvals to prevent unauthorised transactions. A WP Intelligence report by Dan Gallagher, Lead AI & Tech Analyst at The Washington Post, says sellers of agentic products must address a trust deficit with consumers and persuade sceptical merchants to grant access to their online marketplaces. It adds that existing regulation on complex transactions and on competition in the tech industry has held back wide adoption, despite recent agentic releases from Meta and OpenAI. [20] [22]

Consumer demand is limited so far. A 2026 Gartner survey found only about 30% of U.S. consumers willing to let AI narrow their purchase choices, and consumers generally prefer AI help with research and price comparison rather than final buying decisions. Against that, McKinsey projects the U.S. business-to-consumer agentic commerce market could reach $1 trillion by 2030, and $3 trillion to $5 trillion globally. For decision-makers, the stories point to access and trust as the constraints: agents may depend on authorised data-sharing routes and on standard identification and permission schemes rather than on customer credentials. [21] [20] [22]

> **Who should take heed.** Bank security and digital leaders, and retailers and platforms building agentic checkout, should take heed: JPMorgan is barring agents that log in with customer credentials, so agent access is likely to depend on authorised data-sharing and permission standards.

GOVERNANCE

Research finds enterprises are not ready to govern agentic AI

TECHNOLOGY HEALTHCARE FINANCIAL SERVICES

OutSystems published research with CIO Dive and KPMG based on a survey of 858 IT leaders. It found that 82% of executives believe their organisations can govern, integrate and monitor AI agents in production. Only 26% feel well prepared to resolve conflicts among agents, systems and workflows. Some 83% oversee agents with existing IT, security or compliance processes, and 5% have automated controls and audit trails. Some 71% expect agentic systems to handle multi-step workflows within five years. The stories do not give a publication date. Both accounts of the report say 74% prioritise integration with existing systems. They differ on the next figure: one gives 63% for agent orchestration, the other 62% for the option to switch models and vendors. [34] [35]

Flashpoint published its Enterprise Threat Intelligence Report: 2026 on 13 October. It found that 25% of cybersecurity teams use agentic AI in defence, 20% run closed pilots, 35% are exploring adoption within 12 to 24 months and 20% have no plans to test it. Teams named these concerns: confidently wrong decisions, data leakage, auditability, compliance, cost and control over workflows. Executives and boards are keener. Over half expect significant budget increases, 35% forecast rises above 10% next year and none expect cuts. At IBM's Think on Tour EMEA 2026, Ana Paula Assis said 80% of executives expect AI to drive revenue growth but only 20% can clearly explain how. [36] [37]

The sources report high pilot failure rates but frame the figures differently. Startup Fortune cites three: a Gartner prediction that over 40% of agentic AI projects will be scrapped by 2027; a finding from MIT's NANDA initiative that 95% of generative AI pilots deliver no measurable P&L return; and a 2026 survey of 600 C-suite leaders in which 42% of companies abandoned most pilots before production. MedCity News says 95% of hospital AI pilots fail to deliver measurable ROI but names no source. Startup Fortune says pilots fail because success is never defined. Its remedy is a written conversion clause with three parts: a quantified metric, a named decision maker with budget authority and a decision deadline. [38] [39]

For decision makers, the sources recommend similar controls. In the OutSystems research, 81% of executives say testing whole workflows and agent interactions is critical when choosing a platform. Experts quoted in the report recommend building runtime controls and policy-as-code in from the start. MedCity News reports that hospital agents work when tasks are narrow, the tool sits inside the patient's EHR chart and humans make the final decision. Executives at IBM and Atos said governance must be designed in. UK insurer Markerstudy used IBM's Bob tool on six million lines of legacy code and kept human oversight. OutSystems and IBM are each promoting their own platforms for agent governance. [38] [34] [37] [35]

> **Who should take heed.** CIOs, CISOs and boards approving agentic AI budgets should take heed: the research shows executive confidence and spending plans running ahead of preparedness for agent conflicts, automated controls and audit trails.

INFRASTRUCTURE

AMD launches Helion, an open rack blueprint for agentic AI, at the OCP Global Summit

TECHNOLOGY

AMD introduced Helion, an open rack-scale blueprint for agentic AI, at the OCP Global Summit in San Jose, California. It is based on an Open Compute Project spec design and builds on AMD's Helios rack architecture. Helion comes in two main forms. The first is a CPU-only rack holding up to 32,768 sixth-generation EPYC cores, aimed at hyperscalers that run agent execution separately from model inference. The second is an all-in-one rack combining EPYC CPUs with Instinct MI350P GPUs, aimed at enterprises. Both liquid-cooled and air-cooled versions are planned. One report says a token router sends agent requests across frontier and open-source models. AMD plans to contribute the specification to OCP. [40] [41] [42] [43] [44]

Partner-built Helion configurations are expected in November 2026. Reports disagree on who the partners are: one names Cisco, Dell, HPE, Lenovo and Supermicro, while AMD's announcement also names Hon Hai, QCT and Wiwynn. Supermicro says it has added Helion to its Data Center Building Block Solutions portfolio, which makes it one of the first partners to build on the blueprint. Its Chief Business Officer, Vik Malyala, said this would help customers deploy agentic AI infrastructure faster. None of the stories give pricing, performance figures or named Helion customers, so buyers cannot yet compare its cost or throughput with alternatives. [40] [42] [43] [44] [45]

Dan McNamara, an AMD senior vice president, said agentic AI involves diverse workloads that need adaptable compute. Yahoo Finance reports that Helion puts AMD in a better position against Nvidia rack designs such as the NVL72. Jordan Klein of Mizuho Securities pointed to AMD's share among cloud hyperscalers but noted growing competition from Nvidia's Vera CPU and other data-centre platforms. On supply, CEO Lisa Su said demand is strong through 2027 and that AMD plans to increase chip supply after capacity expansions in 2026. AMD is working with Foxconn, TSMC and memory suppliers on this. Meta Platforms is cited as deploying EPYC processors; no story names it as a Helion buyer. [40] [41]

AMD recently beat analyst expectations, reporting a 245% rise in earnings per share and 50% revenue growth to $11.5 billion. Nvidia reported $96.2 billion in sales last quarter. AMD shares have risen about 182% over the past year, compared with about 25% for Nvidia, but fell 2.18% to $594.50 on Monday. For infrastructure buyers, three points stand out: an open, OCP-based design that several server makers plan to build, a November 2026 availability target for partner racks, and a supplier reporting strong demand through 2027. [40] [41]

> **Who should take heed.** CIOs and data-centre infrastructure leaders planning agentic AI capacity for 2027 should take note, because partner Helion racks are due in November 2026 and AMD reports strong demand through 2027.

DEPLOYMENTS

Kyndryl deploys its Agentic AI Framework in Toyota Finance IT operations

FINANCIAL SERVICES TECHNOLOGY

Kyndryl, Toyota Systems Corporation and Toyota Finance Corporation announced a collaboration to change how Toyota Finance runs its IT service operations, using the Kyndryl Agentic AI Framework. The AI agents work on infrastructure design, change management and impact analysis. They draw on customer-specific rules and knowledge, including IT environments, runbooks and operational schedules, to find inconsistencies and risks. Story [33] adds that the

system uses graph-based orchestration integrated with Kyndryl Bridge, with the stated aim of outcomes that can be explained and audited. Toyota Finance will deploy the system in a production environment. [4] [5] [6]

The partners list the expected benefits as faster response to business needs, support for governance standards, better business continuity, higher engineering productivity, consistent review quality and quicker decisions. Toyota Systems described its approach as human-centred AI inspired by jidoka, or "automation with a human touch". Kyndryl said its infrastructure-first approach allows scalable, secure deployment in mission-critical systems. The partners plan to extend agentic AI across the whole IT infrastructure lifecycle. The announcement gave no outcome metrics, no timeline and no contract value. Story [33] notes that all of the benefits are reported by the vendors and that Kyndryl has not published adoption or ROI data. [4] [5] [6]

The announcement sits within a wider pattern. Story [33] reports that between October 13 and 14, 2026, three IT services firms named production customers for agentic AI. DXC reported a 95% average reduction in resolution time and at least 92% diagnostic accuracy across 500 use cases, but named no customers and had no independent validation. HCLTech and Anthropic announced a managed Claude deployment for 25,000 engineers and disclosed no outcome data. Story [33] also cites a Brief Script audit of 150 deployments at 119 organisations: only 22 disclosed cost or financial impact data. [5]

For investors, story [87] treats Toyota Finance, together with a separate WPP agreement, as an operational test of Kyndryl's AI strategy that will affect its margin ambitions and investor confidence. It cites analyst forecasts for 2029 of about $15.2 billion in revenue, which is roughly flat, and earnings rising from $198 million to about $472.5 million. More optimistic forecasts put 2029 revenue at $15.4 billion and earnings at $560.1 million. Fair value estimates point to roughly 7% upside in the share price, with some models suggesting up to 50%. Story [87] names contract renewals and the move away from legacy agreements as risks. [7]

> **Who should take heed.** IT leaders at regulated financial firms who are evaluating agentic AI for infrastructure operations should take note: Toyota Finance is a named production deployment, but no outcome data has been published yet, so they should ask any vendor for measured results before committing.

ECONOMICS

Dow cites Workday Flex Credits as a barrier as the cost of agentic AI comes into focus

ENERGY TECHNOLOGY

Dow, a Workday customer since 2019, has not yet switched on agentic AI. Speaking at Workday Rising in Las Vegas, Sarah Jo Rowley, HR Digital Product & AI Leader at Dow, said Workday's Flex Credit payment model is the main reason. Flex Credits charge by AI consumption rather than by user licence. At Dow, IT usually pays for SaaS services, so the model raises the question of whether HR or IT funds the agents and governs their use. The story does not give the event date, Flex Credit prices or the size of Dow's deployment. [1]

Rowley listed specific concerns: how to monitor usage, how to handle overages, and the loss of employee skills as agents take on tasks staff used to do. She said affordability and accountability for what agents do make governance harder. Dow is also cautious because the US has no federal AI guardrails, which could complicate compliance from state to state. Rowley still praised recent AI progress. She described uses such as guiding employees through performance reviews and alerting managers to unusual attrition, which she said could cut administrative time and give better data for decisions. [1]

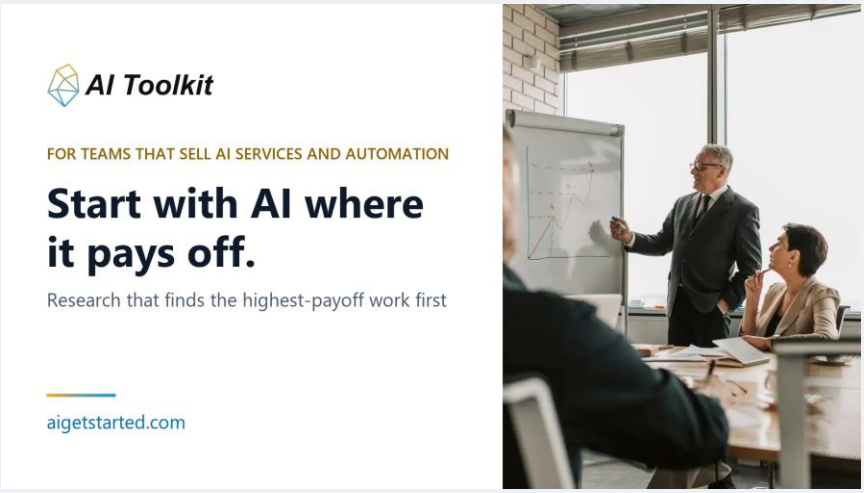
Two other pieces explain why consumption pricing is hard to budget. InfoWorld reports that an agent's work involves several steps, including planning, tool use, checking and retrying. These steps improve accuracy but add latency and cost, so speed, accuracy and cost can rarely all be optimised at once. In Forbes, Pritam Debnath, Global Head of AI & Data Platforms at Sysco, makes the same point. He says the main enterprise challenge is running agentic AI economically at scale, and that the measure that matters is the cost of each successful business outcome, not the cost of the model. [2] [3]

For decision makers, the sources point to a few practical steps. Debnath recommends making consumption explainable through detailed observability, giving business units visibility of their own AI costs while platform teams keep governance,

and using full reasoning only where it adds value. InfoWorld advises teams to agree on acceptable trade-offs early and to keep expert human review in place. Neither piece mentions Workday or Flex Credits directly, so whether these practices would resolve Dow's specific funding question is not covered in the sources. [3] [2] [1]

Who should take heed. CIOs, CHROs and finance leaders weighing consumption-priced agentic AI should take heed, because Dow's experience shows that deciding who funds and governs AI usage can hold up adoption even when the business wants the technology.

ADVERTISEMENT



Know where AI pays off before the first meeting

AI Toolkit is a pre-sales research platform for teams that sell AI services and automation. Enter a prospect's website and it shows how AI solves their problems and which work to start with, then turns that into a live demo, a written report and a pitch deck in your client's brand.

aigetstarted.com

HARDWARE

ASUS, Intel and Netlist launch CPU and memory hardware for agentic AI workloads

TECHNOLOGY FINANCIAL SERVICES

At the OCP Global Summit 2026, ASUS unveiled the XA P2N-E2, a 2U server based on NVIDIA MGX architecture and powered by NVIDIA Vera CPUs. ASUS says it is aimed at agentic AI and reinforcement learning. It supports up to two dual-slot NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs. ASUS also listed rack-scale systems, including the liquid-cooled XA VR721-E3 AI POD with 36 Vera CPUs and 72 Rubin GPUs. It also listed Intel Xeon 6 servers (RS700, RS720) and AMD 6th Gen EPYC 9006 and Instinct systems. The story gives no pricing or availability dates. [8]

Intel used the same summit to introduce Xeon-based rack designs that largely conform to the Open Rack v3 standard. Intel says selected Xeon 6+ configurations deliver over 46,000 cores at about 133.45 kilowatts per rack. Intel also unveiled the GPU D-Series, codenamed Crescent Island: an air-cooled PCIe card rated at up to 1.3 petaflops FP8 within 350 watts, with 160 GB of LPDDR5X as standard and up to 480 GB from partner ODMs. Intel says it targets inference on models up to 300 billion parameters on one card. Customer sampling is underway. No general availability date is given. [9]

Netlist announced DDR5-6000 and DDR5-6400 LIGHTNING Ultra-Low Latency RDIMMs for agentic AI workstations and servers. The modules are JEDEC-compatible and work with AMD and Intel server

platforms. Netlist's internal testing showed up to 10% lower loaded memory latency than standard DDR5 RDIMMs. That figure comes from the company and has not been independently verified. Named target uses include multi-agent frameworks, retrieval-augmented generation, vector databases and high-frequency trading. Netlist says the modules are available now for qualification and deployment, and customers can request evaluation samples from its sales team. The story gives no pricing. [10]

For buyers, the main point is that vendors now pitch CPUs and memory, not only GPUs, as limits on agentic AI. Arm says agent usage on OpenRouter has risen 14-fold since February 2026. It argues CPU performance should be measured by concurrent task completion. Arm cites Signal65 verification showing its AGI CPU does nearly double the modeled rack-level work of AMD EPYC 9755. The stories contain no AMD response. Most performance figures here come from the vendors themselves, so buyers should validate them against their own workloads before committing. [11] [9] [10]

> **Who should take heed.** Data center and infrastructure leaders planning agentic AI capacity, and trading firms running latency-sensitive systems, should take heed: vendors now pitch CPU, rack and memory choices as performance limits, and most of the cited gains are vendor-reported.

IT OPERATIONS

SolarWinds and Plume launch agentic AI products for IT and network operations

TECHNOLOGY MEDIA FINANCIAL SERVICES

SolarWinds announced Avior, a standalone agentic AI system for enterprise IT. It became available on 14 October 2026 as software-as-a-service or self-hosted. It is built on the SolarWinds Platform. It groups related events into single alerts, audits compliance, fixes configuration drift, backs up configurations, manages capacity and costs, and retires ghost assets. Users can set its autonomy anywhere from conservative, where a human must approve actions, to fully autonomous. Every action is recorded and gated by risk. SolarWinds says customer data is encrypted and is not used to train models, and that customers do not need to move data or systems. MSSP Alert reports that managed service providers are expected to offer Avior to their customers. Neither story gives pricing. [26] [27]

Plume Design launched Plume Console, a central tool for internet service providers (ISPs). ISP teams can use it to investigate subscriber issues and to deploy or build agents. It works across Plume WiFi, Security and Care, and connects to ticketing and CRM systems. Plume says it serves over 450 ISPs in more than 60 countries. It says it draws on data from over 50 million homes, 600 million connected devices and 3000 applications. Plume reports that users of its technology see 31% lower churn, 20% fewer technician visits and a 16% improvement in Net Promoter Score. The story does not say how these figures were measured. It does not give a launch date or pricing. [28]

The stories describe the technology in different ways. SolarWinds calls Avior deterministic. Bob Friday, Chief AI Officer at HPE, said in a 2026 presentation that enterprises are moving from deterministic scripts to non-deterministic agentic automation. He said trust, built through repeated successful autonomous actions, is what drives adoption. He forecast substantial AI integration in networks within two years, with network engineers moving towards defining policy and overseeing AI. A separate article, sponsored by BigPanda, cites a 2024 U.S. airline outage that cost $380 million in lost revenue and $170 million in additional expenses. It recommends adopting AI agents step by step, under governance, alongside human teams. [26] [29] [30]

For buyers, the common thread is autonomy that increases in stages under human control. SolarWinds offers risk-tiered autonomy settings. Plume says it balances automation with human oversight. HPE varies its autonomous actions by risk level. Every performance figure in these stories comes from a vendor or from sponsored content, and none reports independent results. Buyers comparing these tools should ask what approval gates apply, how actions are audited, and how the vendor figures were measured. [28] [26] [27] [29] [30]

> **Who should take heed.** IT operations leaders, ISP network and customer-care executives, and managed service providers should take note: SolarWinds and Plume are now selling agentic tools that act on their own, and the only performance figures available come from the vendors.

SERVICES

Accenture and Quantiphi expand agentic AI partnerships; NTT DATA Business Solutions launches Elevate

FINANCIAL SERVICES TECHNOLOGY HEALTHCARE

Accenture and Dell Technologies have expanded their partnership of more than 20 years by creating the Accenture Dell Business Group to help organisations scale AI and modernise infrastructure. The group focuses on running agentic AI workloads on private infrastructure, including through sovereign and neocloud providers, and supports open source and open-weight models. Accenture plans to train more than 3,000 practitioners across Dell's portfolio, with a focus on highly regulated sectors. The story says UK-based AI provider RHAELM used the partnership to speed up its sovereign-grade AI solutions. The story gives no date, no financial terms and no named clients other than RHAELM. [23]

Quantiphi, which describes itself as an AI-native services company, and LangChain announced a strategic partnership. It combines LangChain's agent engineering platform, including LangSmith, LangGraph and Deep Agents, with Quantiphi's forward-deployed engineering. The partners name three focus areas: governance and observability; embedded engineering support; and feedback loops that turn production data and expert input into better agent performance. They say the approach works with both frontier and open-weight models and on several cloud infrastructures. Quantiphi's Siddharth Kotwal and LangChain's Darren Moffett both said the aim is to move customers from prototypes to production. The story gives no date, commercial terms or named customers. [24]

NTT DATA Business Solutions launched Elevate, an AI-powered delivery framework. It combines the company's NTT DATA aXet platform, Natuvion's data transformation expertise and SAP technologies, including SAP's Agent-Led Toolchain and SAP Business AI Platform. Elevate is aimed at SAP lifecycle consulting and ERP modernisation such as RISE with SAP. It comes with AI readiness assessments, governance, adoption services and workforce enablement programmes. The company is also developing GLASSWING for master data management and MAKIRA for document and content management. The story reports no launch date, pricing or client results. [25]

All three announcements pair vendor technology with service-provider engineering, and all stress control: over data, over infrastructure or over how delivery is governed. Two of them, Accenture with Dell and Quantiphi with LangChain, explicitly support open-weight models. For buyers, the stories describe offerings and stated intentions, not results. None of them reports performance figures, costs or completed deployments. The claims about moving AI from pilot to production are therefore not supported by evidence in these sources. [23] [25] [24]

> **Who should take heed.** Technology and procurement leaders in regulated sectors who are weighing private or governed agentic AI deployments should take note, because all three providers are pitching control and production readiness without yet publishing costs or results.

OPEN SOURCE

Perplexity and Google open-source agent tools as the Agentic AI Foundation changes its rules

TECHNOLOGY

The foundation that governs the Model Context Protocol (MCP) and Agent2Agent (A2A) has changed how it admits projects. The Linux Foundation hosts it. The story gives it two names, the Agentic AI Foundation and the AI Agent Interoperability Foundation, and both shorten to AAIF. Over two years it has grown to more than 270 members and seven working groups. Its first entry criteria copied the Cloud Native Computing Foundation and required broad production use, which most early agent projects could not show. It now has three stages: sandbox, growth and impact. Sandbox projects need active maintainers, are reviewed from time to time, can be archived if inactive and get no funding. The story does not say when the change took effect. [31]

AAIF has also opened its working groups to non-members. Members can still propose groups and provide resources. A new lifecycle process lets groups be closed in good time. The foundation says its focus is now security, identity, authentication and deployment patterns for agentic workloads, which the story calls areas without widely adopted standards. It

does not endorse specific solutions. For teams building on MCP or A2A, this means early-stage projects face a lower bar to join the foundation, and people can contribute to working groups without becoming members. All the governance changes and discussions are documented publicly. [31]

Perplexity has open-sourced SPACE, the sandbox Perplexity Computer uses to run agent tasks in isolated Firecracker microVMs. It is written in Rust (release 0.8.0), licensed under Apache 2.0, and needs Linux x86_64 hosts with KVM. Credentials never enter the guest, and host-side gateways control outbound traffic. Perplexity reports that median VM creation latency in production fell from 185 ms to 60 ms. In a red-team test across 9 models, none of 108 attempts reached host-side secrets. In 11 of 54 runs there was a partial network bypass, closed once source-address validation and TLS termination were added. The story attributes these results to Perplexity and does not mention any independent verification. [32]

Google Cloud AI Research has open-sourced Harness^2. It improves what surrounds a frozen model, such as prompts, tools, memory, skills and orchestration, without changing model weights and without labels or test cases. It edits eight components stored as files, judges each edit by comparing how the agent behaves, and keeps lessons in an EXPERIENCE.md file. On WorkBuddy-Bench, GLM-5 with OpenCode gained 11.5 points after one refinement, and weaker models gained more. It runs on Linux with Python 3.12 or 3.13, with inference through Vertex AI. Neither release story gives a date, and neither says SPACE or Harness^2 is an AAIF project. [33] [32] [31]

> **Who should take heed.** Engineering and platform leaders building agents on MCP or A2A should take note: AAIF now makes it easier to bring early projects in and join working groups, and Perplexity and Google have released self-hostable sandboxing and agent-tuning tools whose performance figures, according to the stories, come from the companies themselves.

DEVICES

New local AI PCs and developer workstations reach market from AMD, Asus and NVIDIA partners

TECHNOLOGY

AMD has launched the Ryzen AI Halo Developer Platform, built on the Ryzen AI Max+ PRO 495 processor. AMD says it supports up to 192GB of LPDDR5x unified memory and can run AI models of 300 billion parameters or more locally. The chip combines up to 16 Zen 5 CPU cores, 40 RDNA 3.5 GPU compute units and an NPU rated at up to 55 TOPS. Developers can choose their own models, frameworks, operating systems and workflows. AMD's pitch is that AI work now needs CPU performance for coordination and enough memory to hold models, applications and data at once, not just fast inference. The story gives no price and no availability date. [17]

NVIDIA has launched RTX Spark, a range of slim Windows laptops that pair an NVIDIA Blackwell RTX GPU with an NVIDIA Grace CPU. NVIDIA says the system offers up to 1 petaflop of AI compute and up to 128GB of unified memory, and runs the full NVIDIA AI platform. Prices start at $2,399. ASUS, Dell, HP, Lenovo, Microsoft and MSI are selling the laptops. NVIDIA held launch events in six cities: Berlin, Beijing, Istanbul, Seoul, San Francisco and Taipei. Chief executive Jensen Huang appeared unannounced at the Berlin event. NVIDIA says it offers developer resources for running local AI on RTX Spark machines. [18]

Asus's Ascent QN10 mini PC uses Qualcomm's Snapdragon X2 Elite chip and comes with 32GB of RAM and a 512GB SSD. Its NPU is rated at 80 TOPS, and Asus claims it is the fastest NPU in any mini PC. PCMag UK disagrees with how that claim lands in practice. It says rival AMD and Intel mini PCs deliver more total AI performance by combining GPU and NPU, reaching 131 TOPS (AMD) and 172 TOPS (Intel). In PCMag's tests the QN10 lost to Apple's Mac mini (2026, M6) on CPU, productivity and gaming, and its graphics were weak. PCMag does not recommend it for general desktop use or AI development. The review gives no price. [19]

Buyers now have local AI hardware from three chip camps, but the vendors measure it differently. AMD quotes memory capacity and model size, NVIDIA quotes petaflops, and Asus quotes NPU TOPS alone. PCMag's review shows that a headline NPU figure can understate or overstate what a machine actually delivers. Of the three, only NVIDIA gives a starting price. AMD gives no date or price, and no independent tests of the AMD or NVIDIA systems appear in these stories. Anyone deciding on hardware should ask vendors for prices and wait for benchmarks on the workloads they plan to run. [17] [19] [18]

Who should take heed. IT and engineering leaders who are budgeting for on-device AI development should take heed, because AMD, NVIDIA and Qualcomm-based systems are now on offer with performance claims that cannot be compared directly, and only one of them has been independently tested in these reports.

TECH CORNER

Helion splits the agent loop from the model: how AMD's rack blueprint is laid out

Written for readers who want the mechanism. The rest of the issue does not depend on it.

AMD's Helion is a reference rack blueprint, not a fixed product. AMD describes it as an open design based on an Open Compute Project (OCP) specification and built on the open-platform approach of its Helios rack-scale solution. Supermicro describes it as ORv3 compliant. Partners can tune CPU core density, CPU performance and power, add GPU acceleration, and choose liquid- or air-cooled versions. AMD unveiled it at the OCP Global Summit in San Jose and plans to contribute the specification to OCP for community development. Its ecosystem list includes Cisco, Dell Technologies, Hon Hai Technology Group, HPE, Lenovo, QCT, Supermicro and Wiwynn. Supermicro has added Helion to its Data Center Building Block Solutions portfolio and says it is one of the first partners to build on the blueprint. On timing, AMD's own announcement says the platform is expected to launch in November 2026; other reports say partner-developed rack configurations are expected to be available in November 2026. [41] [42] [43] [44] [45]

The design choice worth noting is the CPU-only configuration. It holds up to 32,768 sixth-generation EPYC cores and is aimed at hyperscalers. AMD positions it as dedicated agent execution capacity that is kept separate from, and connects to, model inference infrastructure. In other words, the work of running agents sits on CPU racks, and the models they call are served elsewhere. The second configuration is an all-in-one enterprise rack. It places EPYC CPU nodes, Instinct MI350P GPU nodes and general-purpose compute on one platform, and AMD says it is intended to run agentic AI alongside existing IT investments. One report adds a token router that directs agent requests across frontier and open-source models. That component appears in only one of the stories, and none of them describes how it works. [40] [42] [43] [44]

AMD's rationale, as stated by senior vice president Dan McNamara, is that agentic AI is made up of diverse workloads that need adaptable compute. The stories do not give performance, power or cost figures for either configuration. Coverage frames Helion against Nvidia's rack designs, including the NVL72. Jordan Klein of Mizuho Securities pointed to AMD's share in cloud hyperscalers and also noted growing competition from Nvidia's Vera CPU and other data-centre platforms. [40] [41]

ADVERTISEMENT



Your Website Should Work Harder Than You Do.

Build your digital presence with The WebSite Guru. Professional, AI-powered websites starting at $500, plus custom e-commerce, chatbots, and lead generation for HVAC, electrical, plumbing, medical, legal, and retail businesses. Get online fast, generate leads smarter, and compete like an enterprise.

www.thewebsiteguru.com

Briefs

Everything else that happened, one line each. These did not cluster into a theme.

Google Cloud Advances Agentic AI and Data Analytics at CaixaBank - The Fast Mode · Banking, Financial Services & Insurance [46]

[Law Firms Now] Kim & Chang to Hold Seminar on Security Issues and Regulatory ... · Technology & Consumer Electronics [47]

Kim & Chang hosts Seoul seminar on agentic AI security and regulation - CHOSUNBIZ · Technology & Consumer Electronics [48]

Artificial intelligence and the future of software - Wilson Asset Management · Technology & Consumer Electronics [49]

How Hong Kong SMEs should set boundaries for agentic AI in B2B selling | Retail Asia · Retail & Ecommerce [50]

The AI alarm is loud. The incident response is too quiet | Lowy Institute · Government & Public Sector [51]

Agentic AI and customer experience: Why autonomy isn't the goal - Foundever · Technology & Consumer Electronics [52]

What's Stopping Your AI Agent From Breaking the Rules You Give It? - HackerNoon · Technology & Consumer Electronics [53]

Don't forget the CX! How Clay puts agents out onto Slack for support but still rates the ... - diginomica · Technology & Consumer Electronics [54]

Google Cloud's New Gemini Agent Rocks Agentic AI Category - Cloud Wars · Technology & Consumer Electronics [55]

Agentic AI in Procurement: High Potential, Higher Organizational Hurdles - Cloud Wars · Technology & Consumer Electronics [56]

The AI-First Tech Function: What CEOs Need to Know · Technology & Consumer Electronics [57]

TD expands agentic AI across mortgage operations to make banking faster, simpler and more human · Banking, Financial Services & Insurance [58]

OPPO Partners with Manus to Bring Agentic AI Experiences to Smartphones - EQS News · Technology & Consumer Electronics [59]

When Agentic AI Moves From Hype to Action - ADWEEK · Media & Communications [60]

From Enterprise to Earthshot: UK's Ecosystem Moves From Ambition to Application With NVIDIA · Healthcare [61]

Agentic AI and the Duty of Care: What the METR Report Means and Does Not Mean for the ... · Banking, Financial Services & Insurance [62]

From Agentic AI to Civil Society Threats: SecTor 2026 Reveals a Rapidly Shifting ... - Yahoo Finance · Technology & Consumer Electronics [63]

Confido Health nabs $12.5M to scale agentic AI across PE-backed medical groups · Healthcare [64]

Druid AI Expands Enterprise Agentic AI Platform with Voice Enhancements, Agent ... - Rutland Herald · Healthcare [65]

Anugal Introduces Agentic AI-Led IGA Modernization Strategy: Replace, Migrate or Extend · Technology & Consumer Electronics [66]

Brazil Firms Modernize Applications with Governed Agentic AI - Company Announcement · Banking, Financial Services & Insurance [67]

Under the Hood: How to Secure Agentic AI on AWS with Check Point - BrightTALK · Technology & Consumer Electronics [68]

Agentic AI doesn't have a model problem. It has a retrieval problem. - The New Stack · Technology & Consumer Electronics [69]

Sources

Every source opens on ainews.davidsoden.com. If the story has rolled out of the feed the site says so and offers the original publisher.

[1] Workday Rising - Dow wants to adopt agentic AI, but Flex Credits are a barrier - diginomica · diginomica.com

[2] 5 agentic AI trade-offs to consider before production - InfoWorld · infoworld.com

[3] Engineering The Economics Of Agentic AI - Forbes · forbes.com

[4] Kyndryl, Toyota Systems and Toyota Finance modernize IT operations with agentic AI · kyndryl.com
[5] Kyndryl and Toyota Finance Deploy Agentic AI Inside IT Operations. The Numbers Can Wait. · finance.yahoo.com
[6] Kyndryl, Toyota Systems and Toyota Finance Modernize IT Operations with Agentic AI · prnewswire.com
[7] Did AI Client Wins Just Shift Kyndryl (KD) Investment Narrative? - Simply Wall St News · simplywall.st
[8] ASUS Unveils XA P2N-E2 Powered by NVIDIA Vera CPU for Agentic AI Workloads at OCP 2026 · press.asus.com
[9] Intel: Xeon Racks and Crescent Island for Agentic AI - Igors Lab · igorslab.de
[10] Netlist Introduces Next-Generation LIGHTNING(tm) ULL DDR5 RDIMMs for Agentic AI · prnewswire.com
[11] AI agents are changing what CPU performance means - Arm Newsroom · newsroom.arm.com
[12] The ICO's agentic AI call for evidence: What you need to know | TLT LLP · tlt.com
[13] Global financial watchdog calls for more human oversight of agentic AI · tribune.com.pk
[14] Singapore Launches World's First Accredited Agentic AI Testing Ecosystem - Tech Times · techtimes.com
[15] ERA Living Guidelines on the responsible use of Generative AI in research updated · research-and-innovation.ec.europa.eu
[16] The Digital Omnibus: Recalibrating Compliance for Agentic AI - Yahoo Finance · finance.yahoo.com
[17] Meet the Next AMD Ryzen(tm) AI Halo · amd.com
[18] RTX Spark Ignites Global Celebration as New Windows PCs Arrive - NVIDIA Blog · blogs.nvidia.com
[19] Asus Ascent QN10 Review: This Climber Gets Grounded by the Mac mini - PCMag UK · uk.pcmag.com
[20] 'You don't want rogue agents': Jamie Dimon bars outside AI agents like Meta's Muse from ... · fortune.com
[21] What AI Shopping Agents Actually Do When They Start Buying Products on Your Behalf · newsobserver.com
[22] AI & Tech Brief: Agentic money's roadblocks - The Washington Post · washingtonpost.com
[23] Accenture, Dell Expand Partnership to Scale Enterprise AI and Modernize Infrastructure · thefastmode.com
[24] Quantiphi and LangChain Partner to Scale Production-Ready Agentic AI for Enterprises · prnewswire.com
[25] NTT DATA Business Solutions launches "Elevate", its Agentic AI-powered delivery ... · prnewswire.com
[26] SolarWinds Unveils Avior, Bringing Deterministic, Agentic AI to Enterprise IT | AFP.com · afp.com
[27] SolarWinds unveils Avior, an agentic AI system for IT management | news | MSSP Alert · msspalert.com
[28] Plume Launches Agentic AI Platform to Help ISPs Automate Network Operations · thefastmode.com
[29] Agentic AI to Prevent Incidents and Cut Resolution Time · emerj.com
[30] Inventor of AI Ops for enterprise Bob Friday explains why agentic AI will finally make self ... · wifinowglobal.com
[31] What the Agentic AI Foundation's rule changes mean for anyone building on MCP or A2A · diginomica.com
[32] Perplexity Open-Sources SPACE, a Firecracker MicroVM Sandbox for AI Agents · marktechpost.com
[33] Google Open-Sources Harness2, a Recursive Agent Harnessing Framework · marktechpost.com
[34] The Agentic AI Era Is Here--Enterprise Readiness Is Not, New OutSystems Research Finds · finance.yahoo.com
[35] The Agentic AI Era Is Here--Enterprise Readiness Is Not, New OutSystems Research Finds · finance.yahoo.com
[36] Defenders Reluctant to Deploy Agentic AI as a Cybersecurity Tool - Infosecurity Magazine · infosecurity-magazine.com
[37] IBM: Agentic AI is reshaping governance, controls and trust - TechTarget · techtarget.com
[38] Why Agentic AI Works in Hospitals When You Strip It Down - MedCity News · medcitynews.com
[39] How to Structure the AI Agent Pilot to Production Handoff - Startup Fortune · startupfortune.com
[40] AMD Makes Big Agentic AI Push With Helion - Advanced Micro Devices (NASDAQ:AMD) · benzinga.com
[41] AMD launches Helion rack server for agentic AI - Yahoo Finance · finance.yahoo.com
[42] Introducing AMD Helion, Agentic AI Compute at Rack Scale · ir.amd.com
[43] AMD Unveils Helion Rack Designed for Agentic AI | TechPowerUp · techpowerup.com
[44] AMD introduces open Helion rack blueprint for agentic AI - Engineering.com · engineering.com
[45] AMD Unveils Helion: Open Rack Blueprint for Agentic AI | Data Centre Magazine · datacentremagazine.com
[46] Google Cloud Advances Agentic AI and Data Analytics at CaixaBank - The Fast Mode · thefastmode.com
[47] [Law Firms Now] Kim & Chang to Hold Seminar on Security Issues and Regulatory ... · asiae.co.kr
[48] Kim & Chang hosts Seoul seminar on agentic AI security and regulation - CHOSUNBIZ · biz.chosun.com

© 2026 David Soden · Get this in your inbox every week: ainews.davidsoden.com/#subscribe

[49] Artificial intelligence and the future of software - Wilson Asset Management · wilsonassetmanagement.com.au
[50] How Hong Kong SMEs should set boundaries for agentic AI in B2B selling | Retail Asia · retailasia.com
[51] The AI alarm is loud. The incident response is too quiet | Lowy Institute · lowyinstitute.org
[52] Agentic AI and customer experience: Why autonomy isn't the goal - Foundever · foundever.com
[53] What's Stopping Your AI Agent From Breaking the Rules You Give It? - HackerNoon · hackernoon.com
[54] Don't forget the CX! How Clay puts agents out onto Slack for support but still rates the ... - diginomica · diginomica.com
[55] Google Cloud's New Gemini Agent Rocks Agentic AI Category - Cloud Wars · cloudwars.com
[56] Agentic AI in Procurement: High Potential, Higher Organizational Hurdles - Cloud Wars · cloudwars.com
[57] The AI-First Tech Function: What CEOs Need to Know · bcg.com
[58] TD expands agentic AI across mortgage operations to make banking faster, simpler and more human · td.mediaroom.com
[59] OPPO Partners with Manus to Bring Agentic AI Experiences to Smartphones - EQS News · eqs-news.com
[60] When Agentic AI Moves From Hype to Action - ADWEEK · adweek.com
[61] From Enterprise to Earthshot: UK's Ecosystem Moves From Ambition to Application With NVIDIA · blogs.nvidia.com
[62] Agentic AI and the Duty of Care: What the METR Report Means and Does Not Mean for the ... · corpgov.law.harvard.edu
[63] From Agentic AI to Civil Society Threats: SecTor 2026 Reveals a Rapidly Shifting ... - Yahoo Finance · finance.yahoo.com
[64] Confido Health nabs $12.5M to scale agentic AI across PE-backed medical groups · fiercehealthcare.com
[65] Druid AI Expands Enterprise Agentic AI Platform with Voice Enhancements, Agent ... - Rutland Herald · rutlandherald.com
[66] Anugal Introduces Agentic AI-Led IGA Modernization Strategy: Replace, Migrate or Extend · prnewswire.com
[67] Brazil Firms Modernize Applications with Governed Agentic AI - Company Announcement · markets.ft.com
[68] Under the Hood: How to Secure Agentic AI on AWS with Check Point - BrightTALK · brighttalk.com
[69] Agentic AI doesn't have a model problem. It has a retrieval problem. - The New Stack · thenewstack.io

© 2026 David Soden · Get this in your inbox every week: ainews.davidsoden.com/#subscribe