

Issue 15, May 2025

Welcome to *The Skinny on AI for Education* newsletter. Discover the latest insights at the intersection of AI and education from Professor Rose Luckin and the EVR Team. From personalised learning to smart classrooms, we decode AI's impact on education. We analyse the news, track developments in AI technology, watch what is happening with regulation and policy and discuss what all of it means for Education. Stay informed, navigate responsibly, and shape the future of learning with *The Skinny*.

Headlines

- [AI is Easy, People are Hard: Culture is the Real Innovation](#)
- [What the Research Says about Self-Directed Learning and AI](#)
- [The 'Skinny Scan' on What is Happening with AI in Education](#)
- **AI News Summary**
 - [AI in Education](#)
 - [AI Ethics and Societal Impact](#)
 - [AI Employment and the Workforce](#)
 - [AI Development and Industry](#)
 - [AI Regulation and Legal Issues](#)
 - [AI Market and Investment](#)
- [Let's also learn from Andrew Ng's 'The Batch'](#)
- **Further reading: find out more from these resources**

Welcome to *The Skinny on AI in Education*. In our **What the Research Says (WTRS)** feature, I bring educators, tech developers and policy makers actionable insights from educational research about metacognition. Fancy a broader view? Our signature **Skinny Scan** takes you on a whistle-stop tour of recent AI developments reshaping education.

But first I wanted to share some thoughts prompted by what I've been doing and reading over this last month...

AI Is Easy, People Are Hard: Culture Is the Real Innovation

The true challenge of the AI era isn't smarter systems — it's building environments where humans can thrive alongside them.

In boardrooms and policy discussions across the country, we earnestly debate the skills needed for an AI future. We reference impressive frameworks and produce comprehensive lists of "21st-century competencies" and the like. But these conversations skip the most essential question: what precisely does success look like? And critically, what kind of learning environments and organisational cultures will nurture these capabilities?

Like all technology-driven transformations before it, success in the AI 'revolution' is about people not machines. Ultimately, success hinges on effective change management with the greatest challenge is not developing more sophisticated AI systems but **nurturing the human capabilities and cultural conditions that will enable people to flourish alongside these technologies.**

Of course, we cannot precisely predict how the AI story will unfold. The contours of future workplaces, the specific applications that will emerge as winners, and exactly how society will evolve remain a mystery. However, this uncertainty should not prevent us from developing a compelling vision of what we want people to achieve in this changing landscape. While specific technical skills may or may not quickly become obsolete, we can articulate the human learning capabilities and cultural conditions that will enable ongoing adaptation, resilience and growth.

A meaningful vision must go beyond cataloguing skills to articulating the cultural and contextual transformation required. We need to envision learning ecosystems where continuous development is woven into the fabric. This means reimagining the very environments where people work and learn. Traditional educational institutions with rigid curricula and assessment systems may fundamentally

hinder the nimble, curiosity-driven learning our AI future demands. Similarly, workplaces that fail to create psychological safety, dedicated learning time, and reward structures that encourage experimentation will stifle the very capabilities they claim to value.

The renowned management principle that "culture eats strategy for breakfast" has never been more relevant. We can devise the most sophisticated strategies for AI education and workforce development, but if they're implemented within cultures resistant to continuous learning, they will inevitably fail.

Many institutions claim to value learning while maintaining structures that actively discourage it. Our vision must confront this contradiction directly. The transformation we need isn't about better training programmes or more educational technology. It requires reconceptualising learning not as something that happens at designated times in designated spaces, but as the constant, natural activity of a healthy organisation or community. In this vision, leaders become learning architects who create the conditions where continuous development thrives because it's embedded in the culture, rewarded in practice, and supported through thoughtful systems design.

[Subscribe to The Skinny](#)

What the Research Says about Self-Directed Learning and AI

Now, let's explore **"What the Research Says" about Self-Directed Learning and AI**

For a longer version of this article with references please

visit: <https://www.educateventures.com/what-the-research-says>

Self-directed learning involves individuals taking initiative for their own learning journey—identifying needs, setting goals, finding resources, and evaluating progress. Self-regulated learning focuses more specifically on the cognitive, metacognitive, and motivational processes within specific learning activities. These learning approaches have become increasingly important in our information-rich, rapidly evolving educational landscape.

The Emergence of AI in Supporting SDL

AI technologies can enhance SDL in several valuable ways. AI enables personalisation by analysing individual learning patterns to provide tailored content and feedback, addressing the challenge of optimising learning for individual needs. These systems offer scaffolding and guidance, providing support when needed and gradually withdrawing as learners develop competence. The technology delivers real-time feedback, offering immediate responses to learner actions, helping them understand progress and adjust strategies accordingly. Furthermore, AI algorithms can guide learners toward information that reduces uncertainty about a topic, supporting more effective knowledge acquisition.

Research findings also highlight important limitations and concerns with AI-supported learning. Over-reliance on technology may lead learners to become dependent on technological scaffolding, potentially reducing active intellectual engagement. Algorithm limitations present another challenge, as incorrectly specified learning models can lead to biased information sampling and ineffective learning experiences.

Several studies emphasise the importance of developing critical AI literacy to prevent overreliance and address the "hallucination effect" where AI generates incorrect information. There's also tension regarding whether AI truly enhances educational outcomes or merely changes processes without substantive improvements. Additionally, research raises questions about whether GenAI might exacerbate existing educational disparities due to unequal access and inherent biases.

Recommendations Based on Evidence

For Educators

1. **Balance facilitation with guidance:** Relinquish some authority while still providing essential structure, especially for younger learners.

2. **Foster AI literacy:** Teach students to critically evaluate AI-generated content and develop validation strategies.
3. **Integrate metacognitive scaffolding:** Incorporate prompts that encourage learners to reflect on their strategies when interacting with AI tools.
4. **Address equity proactively:** Recognise and mitigate potential barriers to equitable access to AI-enhanced learning opportunities.

For Technology Developers

1. **Design for comprehensive SDL cycles:** Support the complete self-directed learning process, not just isolated components.
2. **Bridge optimisation gaps:** Create interfaces that help learners fully leverage AI capabilities.
3. **Balance automation with agency:** Ensure AI provides suggestions rather than prescriptive pathways, preserving learner decision-making.
4. **Address diversity:** Design tools that adapt to various cultural contexts and learning populations.

For Policy Makers and Educational Leaders

1. **Recognise SDL as core:** Position SDL as an essential 21st-century skill in curriculum frameworks.
2. **Invest in teacher development:** Provide substantial professional development focused on technology-enhanced SDL.
3. **Support research:** Fund longitudinal studies examining long-term impacts of AI-enhanced SDL.
4. **Develop ethical frameworks:** Create comprehensive guidelines addressing privacy, bias, and responsible AI use.

By grounding educational practices, technology design, and policy in solid research on SDL and SRL, we can create environments that empower learners while preparing them for lifelong learning in an evolving digital landscape. Evidence-based approaches are crucial to ensuring that AI enhances rather than diminishes human learning capabilities, creating a future where technology amplifies human potential.

The 'Skinny Scan' on what is happening with AI in Education....

My take on the news this month – more details as always available below:

In classrooms, lecture halls, and educational institutions across the globe, AI is slowly starting to take seed. Across the public sector too, there are indications of AI, is increasing uptake. From the BBC's effort to personalise public broadcasting with AI, to the NHS's ongoing struggle to implement digital patient records effectively, the education and public sectors are grappling with how to turn potential into practice. Higher education is engaging more, and more, although often in quite discreet ways, rather than seeing the potential for systemic transformation. And students remain caught in a gap—enthusiastic about AI's promise, but often undertrained and unsupported. Meanwhile, educators are being encouraged to consider the systemic, cognitive, and pedagogical implications of AI to avoid a future where AI supports convenience at the cost of critical thinking. The narrative is clear: the tools are here, but the culture, training, and infrastructure still lags behind.

In the workplace, AI is stepping into spaces that were once considered uniquely human—ethics, recognition, communication, and even legislation. In a world-first, the UAE is entrusting AI with helping to write its laws, a move both lauded as visionary and criticised as dangerously unaccountable. Companies like Workhuman are redesigning workplace praise with AI-generated authenticity, while scientists work to restore speech via brainwave-reading implants, pushing the frontier between human cognition and machine interpretation. The emotional and ethical implications of these developments are vast. As AI becomes more embedded in human interactions, questions of dignity, meaning, and trust are not just philosophical—they're urgent design challenges.

Coding, once the domain of highly trained specialists, is increasingly being handled by generative tools that can write, debug, and optimise code from natural language instructions. Yet, this efficiency boost isn't being shared equally. A study from Denmark warns that the unequal adoption of tools like

ChatGPT may be reinforcing existing socioeconomic divides, as those with higher education and digital literacy surge ahead. At the same time, the daily use of AI is becoming more grounded—supporting creativity, simplifying tasks, and augmenting rather than replacing expertise. AI is not eliminating jobs en masse, but it is transforming what many of them look like.

We are also seeing a great deal, more real-world deployment of AI, it's now catching shoplifters, managing workflows in hiring firms, and powering multilingual chatbots on mobile phones. In Europe and Asia, small and mid-sized companies are finding practical applications for AI that deliver results—reducing retail theft, detecting fraud, and even replacing administrative slog. Meanwhile, research powerhouses like DeepMind are playing a tighter hand, slowing down publications to guard their competitive edge.

Across the globe, AI is a big part of the news agenda. China's rise in humanoid robotics and mobile-compatible AI models shows a strong push toward accessible, embedded AI, while India pursues cost-effective, problem-solving models that suit its linguistic and economic diversity. From photonic circuits in Japan to frugal innovation in India, the AI industry is rapidly diversifying in both focus and geography.

And of course, the vast amount of money being pumped into AI development and infrastructure continues apparently unfettered. OpenAI's \$300 billion valuation and mammoth SoftBank-led funding round signal enduring investor confidence in foundation models, even as start-ups like CoreWeave face painful IPO corrections. The race to commercialize AI applications is red-hot, with nimble start-ups hitting \$200 million annual revenues in record time, fuelled by demand for practical tools that improve search, legal services, and coding. Yet, the glow of AI hype is being tempered by geopolitical shocks—new US chip export controls have wiped billions off Nvidia's balance sheet, while global trade wars threaten to splinter supply chains. Investors are learning to distinguish between speculative heat and sustainable growth, even as start-ups draw eye-watering valuations without public-facing products.

AI Trends Summary – May 2025

Before launching into the detailed news section, it's worth taking a moment to look at what other people working in this area I have been saying over the last month.

Andrew Ng's Vision of AI: A New Digital Dawn <https://www.deeplearning.ai/the-batch/>

In his recent newsletters, Andrew Ng has offered some fascinating insights. He paints a picture of AI not as a replacement for human capability, but as a powerful assistant that's reshaping how we work. He shares how AI has transformed his own work, allowing him to venture beyond his Python comfort zone into front-end development—a bit like having a talented translator at your side when visiting a foreign country.

The AI landscape, according to Andrew, is evolving rapidly but thoughtfully. He suggests we needn't fear a future where technical knowledge becomes obsolete. Rather, he emphasises that understanding fundamental concepts remains valuable—like knowing basic grammar helps when using translation tools, understanding core computing principles helps when working with AI assistants.

Andrew is particularly excited about how AI tools are becoming more accessible. He notes that smaller, specialised AI models can now be tailored with relatively few examples, making powerful technology available to smaller organisations and projects. It's rather like how mobile phones evolved from expensive novelties to essential tools available to almost everyone.

Education stands out as an area where Andrew sees particular promise. Rather than simply replacing teachers, he highlights how AI can support them—offering suggestions based on proven teaching methods and helping less experienced educators develop their skills.

George Siemens on AI in Higher Education - <https://buttondown.com/SAIL/>

George Siemens, in his "Sensemaking, AI, and Learning" (SAIL) newsletter, presents several key observations about artificial intelligence's impact on higher education:

Siemens argues that higher education institutions must undergo fundamental transformation as AI increasingly excels at the discrete knowledge tasks that have traditionally formed the basis of modern education. He describes this as a "cognitive escalation" where previously university-taught skills are now automatable. Universities must pivot from "teaching learners certain types of things" to "teaching learners to be certain types of ways" - essentially a shift from knowledge acquisition to developing ways of being and thinking. Siemens also expresses disappointment that many traditional university innovators are focusing on simple point-based AI solutions (like basic chatbots) rather than embracing systemic change. He notes that student AI use varies significantly across disciplines.

He raises concerns about:

- AI-driven unemployment potentially rendering entire industries obsolete
- Negative effects on core human connection, deep thinking and mental health
- The necessity to approach AI project management differently from traditional software development

- Professor Rose Luckin, May 2025
AI News Summary

AI in Education

BBC to use AI as it expands iPlayer offerings (31 March 2025)

The BBC has announced plans to implement artificial intelligence across its services, including iPlayer, BBC Sounds, World Service, and sports reporting. The broadcaster will offer increasingly personalised services through its digital platforms, similar to how rival streaming services use algorithms to recommend content. AI tools will also be used internally for translating World Service content to develop services in new languages and creating live text pages from multiple football broadcasts. Unlike private sector rivals focused on revenue, the BBC intends to balance AI-driven personalisation with public service editorial aims to avoid pushing viewers into a "narrow path of popular shows." With households paying the license fee declining by 1%, the broadcaster faces financial challenges alongside its technological transformation, expecting a £33 million deficit while planning to save £300 million annually by 2027-28.

Many NHS staff struggle to use electronic records effectively, report finds (9 April 2025)

Despite billions spent rolling out electronic patient records (EPRs) across the NHS in England, a Health Foundation think-tank report has found that many frontline staff cannot use them effectively. While 90% of NHS trusts have EPR systems, a significant number struggle to use them to their full potential. Health Secretary Wes Streeting has identified the transition from "analogue to digital" as one of three major shifts needed to improve care delivery. Key challenges include insufficient training, fragmented funding with trusts procuring their own systems, and serious interoperability issues between different systems. These problems have led to concerning incidents including mislabelled blood samples and misidentified patients. In response, the Health Foundation recommends developing a new strategy "to get EPRs working properly," with Malte Gerhold emphasizing that "the biggest opportunities for improving NHS productivity over the next few years will probably come from getting more out of tech that is already in the system."

Generative AI and the Student Experience: EDUCAUSE 2025 Technology Report (14 April 2025)

The 2025 EDUCAUSE report highlights how generative AI and digital tools are reshaping student experiences and expectations in higher education. While 69% of students report satisfaction with institutional tech support, many view their institutions as lagging in innovative tech adoption. Use of AI in coursework remains cautious—43% of students report no use, often due to restrictive guidance or fears of misconduct. Still, students who anticipate using AI in their careers are more likely to engage with it for tasks like brainstorming, organizing, and content generation. Despite 55% of students expecting to use generative AI professionally, only 20% report receiving relevant training. Meanwhile, the majority still prioritize soft skills such as communication and problem-solving over technical competencies. Preferences are shifting back toward in-person instruction, especially among younger students, though flexible modalities remain important. Challenges persist in accessibility and mental health support, with decreasing satisfaction levels reported, suggesting institutions must improve

inclusivity and holistic student services.

Mapping a Multidimensional Framework for GenAI in Education (2 April 2025)

A new EDUCAUSE report outlines a multidimensional framework to help higher education institutions integrate generative AI thoughtfully and responsibly. The framework proposes four dimensions: definitional, systemic, cognitive processing, and pedagogical. These perspectives guide educators in understanding what GenAI is, how it functions, its immediate and long-term impacts, and the cognitive processes students need to engage critically with the technology. Concerns include risks around academic integrity, student over-reliance on AI, and broader societal shifts in cognition and learning outcomes. The systemic view emphasizes how AI's structural, functional, and interactional effects ripple across institutions and society. Cognitive development, particularly fostering epistemic cognition—or critical thinking about knowledge itself—is seen as essential. The article calls for careful, holistic, and critically framed adoption of GenAI tools to enhance education while safeguarding human intellectual development

AI Ethics and Societal Impact

AI praise-giving tool promises 'authentic' insights (13 April 2025)

Workhuman, an Irish tech company with \$1.2bn revenue, has introduced an AI upgrade to its "social recognition" platform, which enables colleagues to post praise for each other's work. The tool, named "Human Intelligence," coaches users to deliver more meaningful feedback while flagging inappropriate language. Despite this technological enhancement, Workhuman emphasizes keeping recognition "human-generated" and "organic." The platform allows staff at companies like BP, Cisco and LinkedIn to redeem praise notes for vouchers or merchandise, with the AI element helping determine appropriate reward levels. For managers, the system provides valuable data insights—from identifying suitable mentors to spotting high-performing staff worth retaining. While workplace culture consultant Bruce Daisley warns that AI risks "taking the humanity out of heartfelt actions," a Harvard Business School study found AI assistance can produce "positive emotional responses" comparable to human teammates. The success of such tools may ultimately depend on organisational culture, potentially being "incredibly helpful" in some environments but merely "another part of performative bureaucracy" in others.

UAE set to use AI to write laws in world first (15 April 2025)

The United Arab Emirates is pioneering an unprecedented plan to use artificial intelligence for drafting and reviewing legislation. Ministers have approved the creation of a new Regulatory Intelligence Office to oversee this initiative, which aims to help write new laws and review existing ones. The government plans to create a massive database of federal and local laws, along with public sector data such as court judgments and government services, expecting the AI to speed up lawmaking by 70 percent. Sheikh Mohammad bin Rashid Al Maktoum, Dubai ruler and UAE vice-president, stated the system "will change how we create laws, making the process faster and more precise." Experts have mixed reactions, with Rony Medaglia of Copenhagen Business School describing the "underlying ambition to basically turn AI into some sort of co-legislator" as "very bold," while others note the autocratic UAE has had an "easier time" implementing sweeping government digitalisation than many democratic nations. Concerns remain about AI's tendency to "hallucinate" and produce content that "makes sense to a machine" but not necessarily for human society.

The race to turn brainwaves into fluent speech (20 April 2025)

Neuroscientists are making significant progress in developing "voice neuroprosthesis" technology that harnesses brainwaves to restore speech abilities using brain implants and artificial intelligence. Edward Chang, a neurosurgeon at UCSF, recently published research in Nature Neuroscience detailing work with a woman with quadriplegia who trained a deep-learning neural network by silently attempting to say sentences using 1,024 different words. The system achieved a median decoding speed of 47.5 words per minute (about one-third the rate of normal conversation) and reduced the lag between brain signals and resultant audio from eight seconds to one second. Companies like Precision Neuroscience are advancing this field with more densely packed electrodes that capture higher resolution brain signals. Precision has worked with 31 patients and recently received regulatory clearance to leave sensors implanted for up to 30 days, with plans to create "the largest repository of high resolution neural data" within a year. The technology focuses solely on the motor cortex and

cannot decode "inner thoughts," achieving about 98% accuracy but still requiring "tens or hundreds of hours" of training data generation.

Inside Interpol's Innovation Lab (25 April 2025)

Interpol's innovation centre in Singapore operates as a global hub where law enforcement officers study criminal strategies and develop technological countermeasures. The facility showcases cutting-edge tools including underwater drones that detect "parasite smuggling" where drugs are attached to ship hulls, robotic "K9" dogs equipped with sensors for dangerous environments, 3D laser scanners that create high-definition virtual crime scenes, and advanced digital forensics equipment. The Singapore complex includes one of Interpol's three global command centres, providing 24-hour monitoring for police in nearly 200 member countries and observing approximately 3.5 million attempted cyber attacks daily. AI has dramatically transformed criminal tactics, particularly in scams where large language models remove telltale signs of fraud while deepfake technology provides convincing "proof" of legitimacy. In response, researchers are developing methods to link untraceable 3D-printed "ghost guns" to specific printers through distinctive production patterns, highlighting the technological arms race between law enforcement and increasingly sophisticated criminal enterprises.

AI Employment and the Workforce

OpenAI and Start-ups Race to Generate Code and Transform Software Industry (18 April 2025)

OpenAI has released new models (GPT-4.1, o3, o4-mini) that benchmark tests suggest are among the best yet for computer programming, alongside Codex CLI, a freely available AI "agent" for coding tasks. OpenAI's CPO Kevin Weil believes "this is the year... that AI becomes better than humans at competitive code forever," as AI systems' ability to solve coding problems has dramatically improved from 4.4% on the SWE-bench test in 2023 to 69.1% in 2025. According to GitHub research, 92% of US-based developers already use AI coding tools. Major tech companies (Anthropic, Google, Meta) and numerous start-ups are focusing on coding as one of the clearest early applications for large language models. These tools generate entire blocks of code from text instructions, identify errors, and attempt to correct them. Coding start-ups have attracted significant investment: Reflection AI (\$130m), Anysphere (\$105m at \$2.5bn valuation), and Poolside (\$500m at \$3bn valuation). Industry experts suggest the role of software engineers is evolving to focus more on understanding requirements, teamwork, and ensuring products meet user needs rather than writing code from scratch.

Unequal Adoption of ChatGPT May Exacerbate Labour Inequality (April 2025)

A study published in PNAS reveals that the adoption of ChatGPT in Denmark is uneven across socioeconomic groups, potentially widening existing labor market disparities. Researchers analyzed survey data linked to national employment records and found that individuals with higher education levels, greater digital literacy, and higher incomes are more likely to use ChatGPT for work-related tasks. In contrast, those in lower-income brackets and with less education are less likely to adopt the technology, missing out on productivity gains and skill development opportunities. The study suggests that without targeted interventions, the proliferation of generative AI tools like ChatGPT could reinforce existing inequalities in the workforce.

How People Are Really Using Gen AI in 2025 - Harvard Business Review (9 April 2025)

A year after early hype, real-world use of generative AI has settled into three dominant categories: personal efficiency, content creation, and knowledge work augmentation. According to Marc Zao-Sanders, most individuals use GenAI tools like ChatGPT and DALL-E3 to save time on tasks such as writing, brainstorming, and summarizing information. Content creation applications include drafting emails, reports, and marketing materials, while knowledge work augmentation involves tasks like decision-support and data analysis. Contrary to early fears, few users rely on GenAI for deep technical or specialist work without human oversight. Business users are increasingly integrating GenAI into workflows to boost productivity rather than replace expertise. However, the author warns that effective use still requires human critical thinking and a clear understanding of AI's limitations.

AI Development and Industry

AI business technology takes on shoplifters and admin drag (27 March 2025)

While some technologies like blockchain and the metaverse struggle to find practical applications, artificial intelligence, cloud computing, and workflow automation are breaking through to mainstream adoption. Paris-based Veesion (ranked 52nd in FT 1000) has developed AI-powered surveillance software that analyzes CCTV footage to catch shoplifters by detecting suspicious body movements,

used by more than 4,000 stores across 25+ countries, and claiming to reduce shoplifting by up to 60%. Unlike facial recognition, it identifies suspicious behaviors without needing to have previously encountered the individual. SourceWhale (ranked 18th) provides subscription-based workflow automation software for recruitment companies, connecting dozens of IT systems and automating daily tasks. Originally a recruitment agency that developed the software as a sideline, it pivoted during the pandemic to become a B2B software company, growing revenue from €137,000 in 2020 to €7.75 million in 2023 without external venture capital. Polish startup Solidstudio (ranked 73rd) develops software for electric vehicle charging networks and is prioritizing AI technology to detect fraud, such as cloned RFID cards that allow drivers to charge multiple vehicles simultaneously.

DeepMind slows down research releases to keep competitive edge in AI race (1 April 2025)

Google's AI arm DeepMind has introduced a tougher vetting process for publishing research, making it harder for its scientists to share their work. Led by Nobel Prize winner Sir Demis Hassabis, the group is most reluctant to share papers revealing innovations that could be exploited by competitors or that cast Google's Gemini AI model in a negative light. This represents a significant shift for DeepMind, which previously prided itself on releasing groundbreaking papers. Among the changes is a six-month embargo before "strategic" papers related to generative AI are released, with researchers often needing to convince several staff members of the merits of publication. DeepMind has become central to Google's AI efforts after the company merged its London-based DeepMind and California-based Brain AI units in 2023. In one incident, DeepMind reportedly stopped the publication of research showing Google's Gemini language model is not as capable or safe as rivals, especially OpenAI's GPT-4. The policy changes have unsettled some staffers, where success has traditionally been measured through appearances in top-tier scientific journals, with some researchers leaving as a result.

China's AI race creates tension at home (1 April 2025)

Chinese artificial intelligence models continue to outpace expectations, with significant recent updates from two major technology groups: DeepSeek released an improved V3 model, while Alibaba launched a Qwen model efficient enough to run on mobile phones. DeepSeek's upgraded model intensifies competition with US rivals such as OpenAI, showing significant improvements in reasoning and coding. The company has made its model publicly available on Hugging Face, enabling developers worldwide to build and test AI agents more easily. Alibaba's latest Qwen model can process images, audio, and video on laptops and mobile phones, envisioning it powering AI agents that provide real-time assistance, such as audio descriptions for visually impaired users. The AI arms race is creating growing tension among Chinese tech companies, with noticeable market reactions. Shares in Baidu, once considered the face of Chinese AI, have significantly underperformed as its models fall short of peers in both capability and adoption. In contrast, Tencent and Alibaba have seen stronger investor confidence, with Alibaba's shares nearly doubling over the past year due to visible momentum in AI deployment and enterprise integration.

China gains dexterous upper hand in humanoid robot tussle with US (9 April 2025)

China's robotics industry, led by Hangzhou-based Unitree, is emerging as a global leader in humanoid robot development. Other prominent Chinese start-ups include AgiBot, Engine AI, Fourier and UBTEch, who have garnered attention through social media demonstrations, including Unitree's 16 H1 robots performing synchronised folk dance during China's spring festival. Industry forecasts predict substantial growth: Goldman Sachs expects the global humanoid market to reach \$205bn by 2035; Bernstein predicts 50 million annual robot sales by 2050; Citibank forecasts 648 million robots by 2040. China's advantage stems from its established electronics and EV supply chain infrastructure, with many humanoid robot components already manufactured for electric vehicles. Chinese robot costs are significantly lower, with Bank of America estimating Tesla's Optimus robot would cost one-third less if built with Chinese components. The Chinese government has designated humanoid robotics as a strategic industry, providing substantial funding and support. Unitree has reduced humanoid robot costs from up to \$1 million to approximately \$100,000 for its programmable H1 model. Current applications remain limited to police patrols, retail entertainment, and initial factory deployment, with technical challenges including energy inefficiency in bipedal movement and extensive programming requirements.

TDK claims optical breakthrough to tackle generative AI's biggest bottleneck (15 April 2025)

Japan's TDK corporation has demonstrated the world's first "spin photo detector," combining optical, electronic, and magnetic elements to achieve response times of 20 picoseconds (20 trillionths of a second). This innovation could process data 10 times faster than current electronics by adapting magnetic heads technology from hard disc drives. Hideaki Fukuzawa, senior manager at TDK, identifies data transfer as "the biggest bottleneck for AI rather than the semiconductor GPU performance," as current systems use slower electrical signals. The shift towards optical technology leverages the faster travel speed of light, with Professor Arata Tsukamoto of Nihon University noting the device "holds remarkable promise" from both scientific and technological perspectives. TDK plans to provide samples to customers by the end of March 2026, with mass production targeted within three to five years. The photonic integrated circuits market is forecast to expand more than tenfold over the next decade to \$54.5bn due to AI's demands, according to IDTechEx. Additional potential applications include smart glasses for AR/VR and high-speed image sensors. Major AI companies are already developing similar technology, with TSMC aiming for production within five years.

Japan hosts unlikely winners of the global AI boom (16 April 2025)

Japanese industrial companies are finding unexpected growth opportunities in the AI boom despite not being at the forefront of AI software development. Companies that make precision motors and fans, once the backbone of Japan's industrial economy, are now finding new growth in the physical infrastructure that makes AI possible. Many of these companies, including Nidec, had previously bet on electric vehicles as a key growth driver, but slowdowns in EV adoption had hit them hard. The extreme heat and power demands of AI infrastructure are fuelling global demand for cooling and power delivery systems, where Japanese companies like Nidec, Sanyo Denki, and Murata Manufacturing are leaders. Nidec's pivot towards AI infrastructure is already delivering results, with the company reporting record operating profit of ¥175.5bn (\$1.2bn) in the nine months to December. While Japan may be lagging in consumer-facing AI development due to strict data privacy laws and language barriers, its industrial engineering groups are proving increasingly valuable in building the physical infrastructure for AI's global expansion, giving these "unassuming companies" a "surprising second act."

India bets on 'frugal innovation' to catch up in global AI race (16 April 2025)

Prime Minister Narendra Modi's government is leveraging India's tradition of "frugal innovation" and its vast tech talent pool to gain ground in the global AI competition. The government seeks to develop cheaper large language models trained specifically on Indian languages and is focusing on building AI "applications" to solve specific practical problems rather than competing directly with frontier research. Modi launched a five-year \$1.2bn AI strategy in 2024 and has made 10,000 graphic processing units available to Indian researchers and start-ups. However, India received just \$179.3mn of the \$43bn global AI investments made in 2024, compared to China's \$3.3bn and the US's \$34.2bn. A 2024 Stanford University report ranked India top in AI skill penetration, but the country accounts for only 0.22% of global AI patents granted between 2014-2022. Sarvam AI, one of India's leading AI start-ups, is developing a new LLM focused on everyday tasks, while Infosys is transforming into an "AI-first company" by upskilling its 340,000+ workforce. IT firm executives suggest India is moving from "labour arbitrage" to "tech arbitrage" as it builds applications on top of leading models rather than creating foundational models from scratch.

Microsoft unveils AI assistant with 'memory' (4 April 2025)

Microsoft has unveiled an upgraded version of its AI assistant that develops a "memory" to remember user preferences and can take actions on their behalf. The personalised "Copilot" was announced during Microsoft's 50th anniversary event and represents the biggest step for the company's consumer AI unit. Mustafa Suleyman, Microsoft's consumer AI chief and co-founder of Google's DeepMind unit, demonstrated several features including the ability to independently book tickets, make reservations, and shop for goods online. The update includes a podcast-generating feature and a new "Vision" feature enabling Copilot to process information from a user's phone camera. Microsoft is using AI to upgrade its Bing search engine as it seeks to compete with Google, which retains a 90% market share in search. The move comes as Microsoft undergoes a strategic overhaul to reduce its OpenAI dependency, despite maintaining a profit-sharing agreement and access to OpenAI's models until at least 2030. Satya Nadella hired Suleyman in March 2024 from Inflection AI, paying \$650 million to license the start-up's technology and hire most of its talent.

Baidu founder highlights 'shrinking' demand for DeepSeek's text-based AI (19 April 2025)

Robin Li, Baidu's founder, claimed at the company's developer conference that demand for text-based AI models like those developed by DeepSeek is "shrinking." Li released two new multimodal models—Ernie 4.5 Turbo and X1 Turbo—and criticized DeepSeek's popular R1 model for having "a higher propensity for misleading 'hallucinations'" and being "slower and more expensive than other domestic offerings." These comments come as Baidu attempts to reestablish itself as an AI leader in China after being forced to pivot its business strategy, dropping its subscription service to its chatbot and making its models freely available as "open source." The company showcased several use cases for its multimodal models, including an AI avatar platform for merchants to create humanlike figures for livestreams and advertising. Baidu also announced a new AI agent application called Xinxiang and revealed it has built a computing cluster of 30,000 Kunlun P800 AI chips. Despite criticizing DeepSeek, Baidu has integrated the rival's models into its Qianfan enterprise platform and its map and search applications, highlighting the complex competitive landscape in China's AI sector.

Large Models Exhibit "White Bear" Cognitive Vulnerability (22 April 2025)

Researchers at Seoul National University have uncovered a critical cognitive flaw in large AI models, such as Stable Diffusion and DALL-E3, likening it to the "white bear phenomenon" — the human tendency to think about a forbidden thought when trying not to. Their study finds that generative models are structurally incapable of properly processing negation, meaning prompts that attempt to exclude an element often inadvertently highlight it instead. This vulnerability can be exploited through prompt-based attacks to generate prohibited content, even in systems with safety filters. Using techniques inspired by cognitive therapy, the researchers proposed prompt engineering methods that mitigate such attacks, improving defense success rates by up to 48%. The findings raise important questions about the reliability of attention-based architectures and highlight the need for models better able to distinguish the presence from the absence of concepts. Despite the proposed defenses, the study concludes that fundamental architectural changes are needed to truly eliminate this vulnerability.

Brands Target AI Chatbots as Users Switch from Google Search (27 April 2025)

Advertising groups and tech start-ups are racing to help brands boost their visibility in results from AI chatbots like OpenAI's ChatGPT and Anthropic's Claude, marking a new era of "search engine optimisation." Companies such as Profound and Brandtech have developed software to monitor how frequently brands appear in AI-powered services. Brands like Ramp (fintech), Indeed (job search), and Chivas Brothers (whisky) have adopted this software to reach millions of users who now use generative AI for online searches. Research from Bain consultancy found 80% of consumers rely on AI-written results for at least 40% of their searches, reducing organic web traffic by up to 25%, with about 60% of searches now ending without users clicking through to another website. Despite these shifts, Google's parent company Alphabet announced its core search and advertising business grew almost 10% to \$50.7bn in Q1 2025. Profound, which raised \$3.5m in seed funding from Khosla Ventures, offers a data analytics platform for brands to track industry-related queries and monitor performance in AI searches. James Cadwallader, co-founder of Profound, described this shift as "a CDs to streaming moment" that threatens Google's search monopoly, while Denis Yarats, co-founder of Perplexity, notes that gaming these AI systems is harder because "the only sort of true strategy is to be as relevant as possible and provide good content."

Anthropic Maps the Inner Workings of Claude 3.5 Using Attribution Graphs (27 March 2025)

Anthropic researchers have unveiled a novel interpretability method called "attribution graphs" to trace the internal reasoning of their language model, Claude 3.5 Haiku. Drawing parallels to biological systems, the team conceptualizes the model's internal components as "features" that interact within circuits to produce outputs. This approach allows for the visualization of how specific inputs influence the model's responses. Key findings include the model's ability to plan rhymes in poetry, perform multi-step reasoning, and distinguish between known and unknown entities. However, the study also highlights limitations, such as the model's potential to generate hallucinations or exhibit biased reasoning. The attribution graphs offer a promising avenue for understanding and improving the transparency of large language models.

Google's DolphinGemma AI Aims to Bridge Human-Dolphin Language Gap (14 April 2025)

Google, in collaboration with Georgia Tech and the Wild Dolphin Project (WDP), has developed DolphinGemma, the first large language model (LLM) designed to interpret dolphin vocalisations. Trained on over 40 years of acoustic data from Atlantic spotted dolphins, the AI can generate dolphin-like sequences, including clicks, whistles, and burst pulses—sounds associated with specific behaviours like socialising or aggression. Researchers plan to use DolphinGemma alongside the Cetacean Hearing Augmentation Telemetry (CHAT) system, which allows divers to play synthesised dolphin-like sounds linked to objects, observing if dolphins mimic these sounds to request items. While this approach offers a novel method to study dolphin communication, experts caution that such mimicry may not equate to true language comprehension, emphasising the need for careful interpretation of the results.

AI Regulation and Legal Issues

EU lawmakers warn against 'dangerous' moves to water down AI rules (25 March 2025)

The European Commission is considering making parts of the EU's landmark artificial intelligence act voluntary rather than compulsory, following intense lobbying from Donald Trump and Big Tech companies. Architects of the EU AI Act have urged Brussels to halt these changes, calling them "dangerous, undemocratic" and warning they create legal uncertainty. The provisions at risk include requirements for AI companies to ensure advanced models don't produce violent or false content, or enable election interference. The debate centres around a "code of practice" being drafted by experts, including Turing-prize winner Yoshua Bengio, to guide implementation for powerful AI models. US tech companies have been highly critical of the EU's AI regulations, with Meta's head of global affairs Joel Kaplan claiming the code would impose "unworkable and technically unfeasible requirements." US Vice President JD Vance publicly criticised "excessive regulation of AI" at France's AI Summit, warning that "AI must remain free from ideological bias." The new Commission, which began its mandate in December, has already withdrawn a planned AI liability directive as part of a broader deregulation push.

Critics of UK's AI copyright proposal must not 'resist change', says minister (24 March 2025)

UK Technology Secretary Peter Kyle has urged opponents of a new artificial intelligence copyright regime not to "resist change" as the government prepares to rule on proposals that have sparked strong opposition from British musicians and filmmakers. Thousands of people in the creative and media industries have protested against the proposed system that would require every company, artist, or author to opt out of their work being incorporated into AI systems by tech companies. Despite the criticism, Kyle stressed he does not want to pit the country's creative industries against technology companies, noting: "We have the third-largest AI market in the world and we have the second-largest creative arts sector in the world. I will not pit one against the other." The Department for Science, Innovation and Technology is currently weighing more than 11,000 responses to its consultation on the future of copyright and AI. The government is also pushing ahead with a 50-part "AI Opportunities Action Plan" and has received more than 200 expressions of interest from local communities to become "AI growth zones," which would streamline regulations to attract investment for new data centres.

Google Broke the Law to Keep Its Advertising Monopoly, a Judge Rules (17 April 2025)

Judge Leonie Brinkema of the U.S. District Court for the Eastern District of Virginia ruled that Google acted illegally to maintain a monopoly in online advertising technology. This marks the second time in a year that a U.S. court found Google had acted illegally to remain dominant, following an August ruling that the company had a monopoly in online search. The Justice Department and states successfully argued that Google's monopoly in ad technology allowed it to charge higher prices and take a larger portion of each sale. Google was found to have broken the law in two of three areas: tools used by online publishers to host ad space and the software that facilitates ad transactions. According to the government, Google has an 87% market share in ad-selling technology, much of it stemming from its \$3.1 billion acquisition of DoubleClick in 2008. The ad tech business generated \$31 billion in 2023, about 10% of Alphabet's overall revenue. The Justice Department has pre-emptively asked the court to force Google to sell parts of its ad technology business, with both sides having seven days to propose a schedule for the next phase of the case. A three-week hearing on potentially breaking up Google in the search monopoly case begins Monday, as part of a wider regulatory push against tech giants by the Justice Department and FTC.

UK's New Online Safety Rules Target Algorithms and Age Checks (April 2025)

Under the UK's new "Children's Codes," websites must alter content algorithms for young users and enforce stronger age checks—or face significant fines. Published by Ofcom under the Online Safety Act, the rules include over 40 measures aimed at protecting children online. Platforms that host adult or harmful content are required to act swiftly in filtering feeds, limiting exposure, and offering support. New rules mandate robust age verification, simplified terms of service, the ability for children to refuse risky group chat invitations, and a designated executive accountable for safety. Critics argue the rules don't go far enough, with some calling them a "bitter pill." However, regulators and government officials maintain the measures represent a pivotal step, especially in managing algorithmic exposure—where harmful content is often passively delivered to children. Enforcement mechanisms include large fines and, in severe cases, court orders to block non-compliant sites in the UK.

AI Experts Share Worry About Misinformation, Not Job Losses (24 April 2025)

A recent Pew Research Center survey reveals a striking contrast between AI experts and the general U.S. public in their concerns about artificial intelligence. While 47% of AI professionals express excitement about AI's increasing role, only 51% of U.S. adults feel similarly, with the rest more concerned. The most common concern shared by both groups is the spread of inaccurate information—cited by 70% of experts and 66% of the public. Experts are especially worried about AI impersonation (78%) and misuse of personal information (71%). However, only 25% of experts express concern over AI-related job losses, compared to 56% of U.S. adults. Similarly, experts are less likely to fear AI disrupting human relationships (37% vs. 57%). Overall, AI experts tend to have a more optimistic long-term view, with 56% expecting a net positive impact from AI in the U.S. over the next 20 years, in contrast to just 17% of the general public.

AI Market and Investment

OpenAI secures \$300bn valuation after \$40bn SoftBank-led funding round (1 April 2025)

OpenAI has raised \$40 billion in new funding, with SoftBank providing 75% of the investment and Microsoft, Coatue Management, Altimeter Capital and Thrive Capital contributing the remaining 25%. The funding values OpenAI at \$300 billion, making it one of the best-funded private start-ups globally, equivalent to the 27th biggest company in the S&P 500 if it were listed. The financing will be delivered in two parts: an initial \$10 billion followed by \$30 billion to be invested by the end of 2025. The investment comes as OpenAI transforms from a complex non-profit structure into a conventional for-profit company, a conversion required to be completed by year-end. SoftBank has the option to reduce its total contribution to \$20 billion if the for-profit conversion doesn't complete by the end of 2025. OpenAI announced plans to release an "open-weights" AI model in the coming months, partly in response to competition from China's DeepSeek and Meta's Llama, representing a shift from OpenAI's traditional subscription-based model towards making some aspects of their technology more accessible. Sam Altman, OpenAI's CEO, stated this investment would help "push the frontier and make AI more useful in everyday life."

CoreWeave raises \$1.5bn in scaled-back IPO as investors' AI enthusiasm cools (27-28 March 2025)

Cloud computing provider CoreWeave has dramatically reduced the size and value of its initial public offering, raising \$1.5 billion compared to its initial target of \$4 billion, signaling wavering investor demand for AI infrastructure. The company sold 37.5 million shares at \$40 each, significantly below its initial plan to sell 49 million shares priced between \$47 and \$55, giving CoreWeave a market value of approximately \$23 billion. Nvidia, which already owns about 6% of CoreWeave, was set to purchase about \$250 million of the shares. The scaled-back IPO represents a dramatic climbdown from initial discussions with bankers that sought to value the company at more than \$35 billion. Recent challenges for CoreWeave include reported violations of several terms of a \$7.6 billion loan last year, triggering "technical defaults," and news that Microsoft, its largest customer, walked away from some commitments. The IPO comes amid broader market challenges, with the Trump administration's aggressive trade agenda hitting tech companies particularly hard. The Philadelphia Semiconductor index has lost 11% this year, while Nvidia has dropped 19%. Alibaba chair Joe Tsai's warning of a potential "bubble" in data centre construction further dampened investor sentiment during CoreWeave's pre-IPO roadshow.

CoreWeave fails IPO 'hair test' (28 March 2025)

The Lex column describes CoreWeave's scaled-back IPO as failing the "hair test" - a metaphorical assessment of how much risk investors are willing to accept with new listings. CoreWeave's business model of providing rental access to high-end AI chips is described as "a cat's cradle of stakeholder interests" with concerning dependencies: it relies heavily on two customers (Microsoft and OpenAI), has one chip supplier (Nvidia), and both Nvidia and OpenAI are also shareholders, with Nvidia taking more stock in the IPO. Despite signing multi-year contracts with customers projecting \$27 billion in revenue through 2030, the company's long-term demand prospects remain uncertain, while interest payments consumed its entire operating profit in 2024. The Lex column views the IPO's "climbdowndown" as a positive sign that investors are becoming more realistic about AI valuations, contrasting with previous exuberance: "When spirits are rowdy, as they have been whenever AI is involved, that adjustment can dwindle to nothing."

Google DeepMind's drug discovery spin-off Isomorphic Labs raises \$600mn (31 March 2025)

Isomorphic Labs, the drug discovery start-up spun out of Google's DeepMind artificial intelligence unit, has raised \$600 million in its first external funding round, as investors bet on the potential for AI to "solve disease." The fundraising for the London-based company was led by Thrive Capital, the New York-based investor that is also among the biggest backers of ChatGPT maker OpenAI. Other participants included GV (Alphabet's venture capital arm) and additional investment from Alphabet itself, which remains the majority shareholder. The fundraising is among the biggest for a UK AI company, providing a financial boost as Sir Demis Hassabis (who received the Nobel Prize for chemistry last October for developing AlphaFold) pushes to have drugs designed by AI in clinical trials by the end of the year. Isomorphic Labs was spun out from DeepMind in 2021 and has partnered with large pharmaceutical companies including Novartis and Eli Lilly. The company's researchers use DeepMind's AlphaFold 3 to predict the structures of genetic code DNA and RNA, as well as ligands, along with proprietary AI models to run predictions and simulate drug development.

AI 'application' start-ups become big businesses in new tech race (14 April 2025)

Multiple AI start-ups creating practical applications based on large language models are rapidly increasing sales and sparking a race to commercialise this cutting-edge technology. According to Lorenzo Chiavarini of Dealroom.co, many of these start-ups are reaching up to \$200 million in annual recurring revenue in less than two years, "faster than having seen before, and often with very nimble teams driving strong investor interest." Start-ups building AI-powered apps have attracted \$8.2 billion of funding in 2024, up 110 per cent compared with the previous year. Notable funding rounds include Perplexity (AI-driven search engine) raising \$500 million in December at a \$9 billion valuation; Harvey (legal AI) raising \$300 million in February; and Anysphere (maker of coding tool Cursor) raising \$105 million at a \$2.5 billion valuation in January. Anysphere is reportedly generating annual recurring revenues of about \$200 million and is receiving interest from investors willing to value it at \$10 billion or more. Coding app companies have been particularly attractive to investors, with Reflection AI, Poolside, Magic and Codeium all raising hundreds of millions of dollars in the past year. These application companies have the advantage of being able to switch underlying AI models whenever cheaper or better ones come along, with Sierra co-founder Bret Taylor noting his company has changed models "at least five or six times in our short history."

Tech Stocks Sink After Nvidia Reveals Hit From US Curbs on Sales to China (16 April 2025)

Tech stocks led a Wall Street sell-off after Nvidia revealed that new US controls on sales to China would wipe \$5.5 billion from its earnings. The Philadelphia Semiconductor index fell 4.1%, taking its loss for the year to more than 24% with all 30 constituent stocks declining. Nvidia was hardest hit, down 6.9%, while other semiconductor companies including Broadcom, AMD and Arm were also affected. The tech-heavy Nasdaq Composite fell 3.1%, while the S&P 500 lost 2.2%. Federal Reserve chair Jay Powell warned that tariffs could put the Fed's inflation and employment goals at risk. The Nasdaq entered bear market territory in early April, marking a decline of more than 20% from its mid-February high. There was a temporary rebound when the White House announced a 90-day pause to "reciprocal" tariffs, excluding those on China. The market downturn comes amid President Trump's aggressive trade policies, including steep "reciprocal" levies on all big US trading partners, which have particularly affected tech stocks. The World Trade Organization warned that Trump's tariffs could drag the world into a recession, with global output at risk of falling as much as 7%.

Nvidia to take \$5.5bn hit as US clamps down on exports of AI chips to China (15-16 April 2025)

Donald Trump's administration is imposing new controls on Nvidia's ability to sell AI chips to China, with Nvidia's H20 chip (already designed to comply with previous Biden-era export controls) now requiring a special licence for Chinese customers. The US Commerce Department confirmed it is issuing new export licensing requirements for the H20, AMD's MI308, and equivalent chips. Nvidia expects to take a \$5.5bn charge in the quarter to April 27 related to H20 chips for "inventory, purchase commitments, and related reserves." Analysts estimate Nvidia will generate about \$17bn in sales to Chinese customers in the current financial year, with Bernstein analysts saying the H20 accounted for about \$12bn of Nvidia's \$17bn revenues in China over the past year. AMD also announced it expects up to \$800mn in charges. Nvidia's shares fell almost 6% in early trading in New York, pulling down the tech-heavy Nasdaq Composite by 1.7%. The controls mark "another escalation in Donald Trump's trade war with Beijing," following additional tariffs of 145% on China earlier in April. Despite reduced performance compared to Nvidia's top GPUs, the H20 has seen solid demand in China due to the shortage of competitive domestic chip suppliers. Analysts at Morgan Stanley noted the scale of the inventory write down "suggests that the company is not optimistic about being granted licenses." Beijing has taken steps to encourage local tech companies to use homegrown chips from companies such as Huawei, and China could potentially freeze out Nvidia's products with new energy efficiency rules. The controls are part of a broader trade context, with the White House press secretary Karoline Leavitt urging China to cut a new trade deal with the US, saying "the ball is in China's court." The Trump administration has also launched a national security probe that could lead to new tariffs on semiconductors, further escalating tensions in the global technology supply chain.

OpenAI co-founder Ilya Sutskever's new venture SSI valued at \$32bn (11 April 2025)

OpenAI co-founder Ilya Sutskever has raised \$2 billion for his artificial intelligence start-up Safe Superintelligence (SSI), valuing the year-old company at \$32 billion despite having no current product. Sutskever left OpenAI in May 2024 after a failed coup attempt against chief executive Sam Altman and launched SSI in June 2024 with Daniel Gross (former head of Apple's AI efforts) and Daniel Levy (AI researcher). The funding round demonstrates investors' continued strong appetite for backing AI start-ups led by prominent researchers, even amid economic uncertainty in the US. Prominent venture capital firms participated in the latest fundraising, with Greenoaks leading the round with \$500 million investment, alongside Lightspeed Venture Partners and Andreessen Horowitz. SSI previously raised \$1 billion at a \$5 billion valuation in September 2024, showing a dramatic increase in valuation over just seven months. The company aims to create AI models that are "dramatically more powerful and more intelligent" than current cutting-edge models from rivals including OpenAI, Anthropic, and Google. SSI has been notably secretive about its technology approach, even with investors. Sources close to the company indicate it is working on "unique ways of developing and scaling AI models" with a focus on surpassing human intelligence.

Alphabet Shares Rise as Google Search Boosts Profits (24 April 2025)

Alphabet reported double-digit increases in first-quarter results, with revenue rising 12% to \$90.2bn and net income jumping 46% to \$34.5bn year-on-year, both beating analyst expectations. Google's core search and advertising business grew almost 10% to \$50.7bn, surpassing estimates of 8-9% growth despite concerns about competition from AI chatbots like ChatGPT, Claude, and Grok. The company claims its AI-generated answers ("AI Overviews") at the top of search results are boosting engagement rather than cannibalising their advertising business, with monetisation rates "approximately the same" as traditional search links. Cloud computing revenue surged 28% to \$12.3bn, reflecting strong demand for AI-related services, though this growth slowed from 30.1% in the previous quarter due to supply constraints as the company works to bring new data centres online. Alphabet shares rose more than 4% in after-market trading, with the company announcing a \$70bn share buyback programme. A one-off \$8bn gain related to shares in an unnamed private company contributed significantly to the increase in net income. Capital expenditure jumped to \$17.2bn in Q1, up from \$12bn last year, as part of planned \$75bn spending in 2025 on AI infrastructure including chips and networking equipment.

Putin's AI blunder is a gift to opponents (2 April 2025)

According to a Russian intelligence defector, Vladimir Putin does not use the internet or own a smartphone, and required his inner circle to use typewriters a decade ago, contributing to Russia

falling behind in artificial intelligence development. Global sanctions have severely limited Russia's ability to develop a domestic AI sector. Sberbank, the state-owned financial services giant, has acquired only 9,000 GPUs since the Ukraine invasion began in 2022, compared to Microsoft's purchase of nearly 500,000 last year. Russia has lost approximately 10% of its tech workforce to emigration since 2022, further hampering its technological development. Russia ranks 31st globally in AI capacity according to Tortoise Media's Global AI Index, behind all major economies and many smaller nations like Portugal, Norway, Ireland, and Luxembourg. Access to advanced semiconductors is becoming increasingly critical for military capabilities, particularly as AI transforms warfare and autonomous weapons systems develop. The author argues that Russia's technological backwardness presents an opportunity for a "deterrence-by-denial" strategy from its opponents, by continuing to restrict access to advanced chips. The article advocates for three approaches: maintaining semiconductor sanctions even if other sanctions are lifted, encouraging further emigration of tech talent from Russia, and communicating Russia's AI disadvantage to regime insiders to potentially foster discontent.

Builder.ai admits past 'problems' while restating revenues (2 April 2025)

Builder.ai, a Microsoft-backed tech start-up that claims to use artificial intelligence to make apps, has restated its 2023 revenues to \$140 million, according to new CEO Manpreet Ratia. The company has appointed BDO to carry out its first group-level audit, covering the years 2023 to 2027. Previously, the company had not engaged an auditor to sign off its accounts at the group level. The restatement comes as Builder.ai attempts to move past the "tumultuous leadership" of its founder, Sachin Dev Duggal, who stepped down as CEO in February but remains on the board with the title of "chief wizard." Ratia, who is also a managing partner at Jungle Ventures (one of Builder.ai's early backers), confirmed a Bloomberg report that the company had lowered its forecasted revenue for the second half of 2024 by 25 percent. The revenue restatement was necessary because the company had worked with "resellers" in the Middle East who failed to fulfill their promised business commitments, making it "very difficult to collect" from these channels. Builder.ai has raised approximately \$450 million since 2016 from investors including SoftBank, Insight Partners, and Qatar's sovereign wealth fund. The company has reduced its global headcount by about 270 people from a total of 770 since Ratia took over.

Let's also learn from Andrew Ng's 'the Batch'

Let's also learn from what Andrew Ng has been collating...

1. OpenAI Launches Cost-Effective GPT-4.1 Models

Summary: OpenAI introduced GPT-4.1, GPT-4.1 mini, and GPT-4.1 nano—models designed to offer high performance in coding and reasoning tasks at reduced costs. These models support up to 1 million tokens of context and are set to replace GPT-4.5 by July 2025. Why it matters: The new models make advanced AI capabilities more accessible and affordable, potentially intensifying competition with other AI developers. [Read more](#)

2. Hugging Face Acquires Pollen Robotics and Launches Reachy 2

Summary: Hugging Face acquired French company Pollen Robotics and introduced Reachy 2, an open-source humanoid robot priced at \$70,000, aimed at research and education in human-robot interaction. Why it matters: This move reflects the growing investment in affordable, open-source robotics, contributing to advancements in AI-driven physical human-machine interaction. [Read more](#)

3. U.S. Imposes New AI Chip Export Controls

Summary: The U.S. government implemented stricter export regulations on advanced AI chips, affecting companies like Nvidia and AMD, and prompting China to accelerate its domestic chip development. Why it matters: These restrictions could disrupt global AI hardware supply chains and may inadvertently strengthen China's self-sufficiency in semiconductor technology.

4. Meta Introduces MILS for Multimodal Capabilities in Text-Only LLMs

Summary: Meta's Multimodal Iterative LLM Solver (MILS) enables text-only language models to process images, video, and audio without additional training by pairing them with multimodal embedding models. Why it matters: MILS offers a cost-effective method to generate high-quality multimodal

training data, advancing zero-shot capabilities and fuelling further development in AI systems. [Read more](#)

5. Google Releases Gemini 2.5 Pro

Summary: Google launched Gemini 2.5 Pro, a model that leads in benchmark performance, utilising reinforcement learning for enhanced reasoning and long-context capabilities. Why it matters: The release demonstrates continued rapid advancement in AI, reinforcing reasoning-focused training as the new standard and countering concerns about stagnation in AI progress. [Read more](#)

6. OpenAI Adopts Model Context Protocol (MCP)

Summary: OpenAI integrated the Model Context Protocol into its Agents SDK and ChatGPT tools, facilitating easier connection of large language models to various tools and data sources. Why it matters: This adoption promotes standardisation and cross-platform collaboration, streamlining the development of agent-based applications across the AI ecosystem. [Read more](#)

7. Sam Altman Reinstated as OpenAI CEO

Summary: After a brief dismissal, Sam Altman was reinstated as CEO of OpenAI, following employee support and board restructuring. Why it matters: The incident underscores the impact of leadership dynamics on the direction and priorities of major AI organisations.

8. Meta Introduces Byte Latent Transformer (BLT)

Summary: Meta's BLT processes raw text at the byte level, improving handling of typos and low-resource languages by eliminating traditional tokenization. Why it matters: BLT enhances model robustness and language versatility, offering practical improvements in real-world language processing scenarios. [Read more](#)

9. Anthropic Reveals Implicit Reasoning in Claude 3.5 Haiku

Summary: Anthropic discovered that its Claude 3.5 Haiku model performs internal reasoning steps without explicit prompting, using transparent architectural tools. Why it matters: This insight enhances understanding of how large language models process information, aiding in the assessment of their true capabilities. [Read more](#)

10. Meta Releases LLaMA 4 Scout and Maverick Models

Summary: Meta introduced LLaMA 4 Scout and Maverick, vision-language models utilizing a mixture-of-experts architecture for efficient processing and high performance. Why it matters: The release demonstrates that open models are approaching parity with proprietary systems, expanding access and fostering innovation in the AI community. [Read more](#)

11. Alibaba Releases Qwen2.5-Omni 7B Multimodal Model

Summary: Alibaba introduced Qwen2.5-Omni 7B, a multimodal AI model capable of processing text, images, audio, and video inputs, delivering state-of-the-art performance across various benchmarks. Why it matters: This model demonstrates that smaller, open-weight multimodal systems can match or surpass larger proprietary models, promoting accessibility and innovation in AI development. [Read more](#)

12. TabPFN: Transformer Model for Tabular Data

Summary: Researchers developed TabPFN, a transformer-based model trained on 100 million synthetic datasets, outperforming traditional decision tree methods in tabular data tasks without additional training. Why it matters: TabPFN bridges the gap between transformers and tabular data, enabling efficient and accurate analysis without the need for extensive retraining, thus broadening the applicability of transformer models. [Read more](#)

13. Kyutai Introduces MoshiVis: Voice-to-Voice with Vision

Summary: Kyutai enhanced its Moshi voice-to-voice model by integrating visual input capabilities, allowing for real-time discussions about images with low latency and natural conversational transitions. Why it matters: MoshiVis exemplifies efficient adaptation of speech-to-speech models to new media types, paving the way for more natural and resource-efficient multimodal AI interactions. [Read more](#)

14. Cloudflare Launches AI Labyrinth to Counter Web Scraping

Summary: Cloudflare introduced AI Labyrinth, a tool that embeds hidden links in webpages, leading unauthorized AI scrapers to decoy pages, thereby wasting their resources and aiding in detection.

Why it matters: As AI models increasingly rely on web data, tools like AI Labyrinth help content providers protect their data from unauthorized use, highlighting the growing tension between AI development and data privacy. [Read more](#)

15. Studies Reveal Emotional Impact of Chatbot Use

Summary: Research from OpenAI and MIT Media Lab indicates that while ChatGPT can reduce feelings of loneliness, heavy users may experience decreased real-world social interactions and increased dependence on the chatbot.

Why it matters: These findings underscore the need for careful consideration of the psychological effects of AI chatbots, especially as they become more integrated into users' daily lives. [Read more](#)

16. Stanford Develops ZeroHSI for 3D Human-Scene Interaction

Summary: Stanford researchers created ZeroHSI, a method to animate 3D human figures interacting with objects in 3D scenes without relying on motion-capture data, using video-based pose estimation.

Why it matters: ZeroHSI reduces the reliance on expensive motion-capture data, enabling more scalable and diverse 3D human-scene interaction modeling, beneficial for various applications in AI and animation. [Read more](#)

17. Google and Mistral Release Compact Vision-Language Models

Summary: Google's Gemma 3 and Mistral's Small 3.1 are new families of multilingual, vision-language models ranging from 1B to 27B parameters, designed for efficient processing on smaller devices.

Why it matters: These models demonstrate that high-performance vision-language capabilities can be achieved in smaller, more accessible models, facilitating broader adoption and deployment.

18. UC Berkeley Introduces Shortcut Models for Faster Image Generation

Summary: Researchers at UC Berkeley developed "shortcut models" that accelerate diffusion model image generation by learning to take larger noise-removal steps without compromising image quality.

Why it matters: This advancement reduces the computational cost and time required for high-quality image generation, making diffusion models more practical for widespread use.

19. Stanford's Tutor CoPilot Enhances Online Tutoring

Summary: Stanford researchers developed Tutor CoPilot, a tool using GPT-4 to assist online tutors by generating hints, explanations, and questions, leading to improved student pass rates.

Why it matters: Tutor CoPilot demonstrates the potential of AI to support educators, particularly less experienced tutors, by providing pedagogical assistance that enhances teaching effectiveness.

20. REPA Accelerates Diffusion Model Training

Summary: Researchers proposed Representation Alignment (REPA), a loss term for transformer-based diffusion models that accelerates learning by aligning model embeddings with those from pretrained models like DINOv2.

Why it matters: REPA enables diffusion models to achieve better performance with fewer training steps, improving efficiency and resource utilisation in model development.