

AI's Recent History

The Skinny on AI for Education, August 2024

Headlines

- [Learn fast and act slow: The wisdom of being curiously cautious](#)
- [What's happening in technology development](#)
- [What's happening in AI and regulation](#)
- [What's happening in AI investment](#)
- [What's happening in AI in Education and Employment](#)
- [What's happening in AI ethics and societal impact](#)
- [What's happening in AI and Cybersecurity](#)
- [What's happening in the AI workforce and labour market](#)
- [Sources/bibliography](#)
- [Items of Interest](#)

Learn fast and act slow: The wisdom of being curiously cautious

This may be the holiday season, but AI is not standing still, a great deal has happened since I wrote my last newsletter and to be honest, I'm finding it hard to put it all together without making this newsletter too long. So, I have tried to provide short summaries and then links to more detail for those who like it. Here is my perspective on what I've seen in the last month as I write this from beautiful Australia.

Last week I had the pleasure of attending EduTech Australia and participating in a panel discussion about AI in Education. During the panel, I suggested that educators should "learn fast but act more slowly" when it comes to AI. This view wasn't universally welcomed. Fellow panellist, Sal Khan, of Khan Academy fame, was quick to challenge my stance on acting slowly. It gave me pause for thought – had I been wrong to emphasise learning over immediate action?

Upon reflection, I stand by my initial assessment. Yes, we've needed to act swiftly in some areas, particularly in addressing the challenges that ChatGPT and the like pose to academic integrity and on being wary about ethics and data privacy. Every organisation does need an AI policy to guide its members. But beyond these immediate concerns, I firmly believe that careful consideration must precede action.

I understand the pressure we're under. Tech developers, having invested enormous sums of money in AI, are understandably keen to see returns. I recently came across a [report on OpenAI's projected expenditure](#) which could amount to \$8.5 billion in 2024, while its annual revenue is more likely to be around \$3.5 - \$4.5 billion, that is a BIG predicted loss. So, it's clear why there's such a push to sell these technologies to us. There is also a pressure to innovate, but this too can backfire as has happened in South Korea where a proposal to introduce [AI-powered digital textbooks in schools](#) has faced significant backlash from parents and academics.

I caution against rushing headlong into AI implementation. While I genuinely believe AI can bring tremendous benefits to education, these benefits will only materialise if we implement the technology carefully and thoughtfully. There is some mind-blowing talk in Silicon Valley and we need to be wary not to get sucked into the hype. For example, I recently read a [fascinating set of blog posts](#) by someone deeply familiar with big tech's AI development. Their optimism about AI's potential to solve world problems was both inspiring and, I must admit, quite concerning. They had answers for every challenge, from chip shortages to power and cooling issues. Yet, I couldn't help but notice that we're already facing predicted [labour shortages in the chip industry](#) – a problem AI can't yet solve, so I think the optimism is slightly "over-egged". This means that whilst we might be seeing some [price wars when it comes to Large Language Models](#) at the moment, I can't see how the long term costs will be reducing

when investor ROI will be a pressing need. And once an educator or student has subscribed to one of these AI tools and integrated it into their practice will they end up feeling trapped into paying too much of a price?

I fully embrace learning fast and learning as much as we can. As educators, we must stay informed about the rapid technology developments and their implications. The tech industry doesn't take holidays, and significant advancements are happening constantly. By keeping abreast of these changes, we stand a better chance of ensuring that we have a voice in how these technologies are integrated into our educational systems. My message to you is this: let's be critical consumers of AI technologies in education. It's essential that we carefully evaluate any AI investments, ensuring that they truly benefit our students and institutions rather than simply acquiescing to the latest tech trend.

In conclusion, I stand by my advice: learn quickly about AI but implement thoughtfully. It may not be an exciting or a 'sexy' message, but I believe it is the right message. By being curiously cautious, we can harness the potential of AI to enhance education, whilst avoiding the pitfalls of hasty adoption. Let's embrace this exciting future, but let's do so wisely to strike a balance between harnessing the potential of AI and ensuring it serves the best interests of society.

Here's a summary of some of the technical developments in AI that occurred during July-August 2024:

1. Anthropic introduced 'Artifacts', a new feature for Claude 3.5 Sonnet, enhancing user interface for tasks like coding and image generation.
2. Researchers developed Gradient Low-Rank Projection (Galore), a method for memory-efficient AI model training that could potentially enable training on consumer-grade hardware.
3. University of Oxford researchers proposed a method to identify Large Language Model (LLM) hallucinations by calculating entropy based on the distribution of generated meanings.
4. OpenAI launched SearchGPT, an experimental AI-powered online search tool.
5. OpenAI released GPT-4o mini, a smaller, cost-effective version of their multimodal GPT-4o model.
6. Microsoft researchers developed VASA-1, an advanced system for generating expressive talking-head videos.
7. Meta released Llama 3.1, a family of open-source language models rivalling top proprietary models in performance.
8. Microsoft researchers introduced AgentInstruct, a framework for producing synthetic data to fine-tune large language models.
9. Black Forest Labs released the Flux.1 family of text-to-image models.
10. Researchers developed TransAgents, a multi-agent workflow system for translating novels from Chinese to English.
11. Google introduced Gemini Live, an AI voice assistant planned for integration into its Android mobile operating system.

If you want more detail about all of this read on...

- [Professor Rose Luckin](#), August 2024

What's happening in technology development

In terms of the technology, we have seen Anthropic introduce 'Artifacts', a new feature for Claude 3.5 Sonnet, enhancing the user interface for working with large language models.

Simultaneously, researchers developed Gradient Low-Rank Projection (Galore), a new method for memory-efficient AI model training, potentially democratising AI model training by making it possible on consumer-grade hardware. As July progressed, researchers at Radboud University evaluated AI model openness, providing guidelines for developers and policymakers. This work highlighted the importance of openness in AI innovation and its growing regulatory implications.

Artificial Analysis introduced the Text to Image Arena leaderboard, ranking text-to-image models based on public judgements, with Midjourney v6 leading in quality but lagging in speed.

Addressing a critical issue in AI reliability, University of Oxford researchers proposed a method to identify Large Language Model (LLM) hallucinations, calculating entropy based on the distribution of generated meanings. This development is crucial for building trust in AI systems, particularly in fields like medicine and law.

In the corporate world, JPMorgan Chase rolled out LLM Suite, a generative AI product, to employees in its asset and wealth management division. The bank pitched this as capable of performing the work of a research analyst, assisting with writing, idea generation, and document summarisation.

The robotics sector, valued at £58 billion, saw significant acceleration in capabilities due to AI advancements. Improved computer vision and spatial reasoning have allowed robots to gain greater autonomy in various environments, with the global market for humanoid robots valued at over £1 billion and growing 20% annually.

Late July saw OpenAI launch SearchGPT, an experimental online search tool challenging Google's dominance in the search market. This move highlighted the ongoing competition in AI-powered search capabilities. Concurrently, Lord Patrick Vallance, the UK's new science minister, emphasised the need for cyber security as a national priority to safeguard important data troves, announcing a cyber security and resilience bill.

OpenAI also released GPT-4o mini, a smaller, cost-effective version of their multimodal GPT-4o model, setting a new standard for smaller, more efficient AI models. Microsoft researchers developed VASA-1, an advanced system for generating expressive talking-head videos, representing a significant advancement in synthetic media generation.

As August began, Meta released Llama 3.1, a family of open-source language models rivalling top proprietary models in performance. This development highlighted the growing potential of open-weights models and the critical importance of data-centric AI. OpenAI continued to push boundaries by testing SearchGPT, an AI-powered search engine combining web crawling with conversational AI capabilities.

Microsoft researchers introduced AgentInstruct, a framework for producing synthetic data to fine-tune large language models, potentially improving model performance across various applications. However, researchers also discovered that ASCII art could bypass safety measures in large language models, underscoring the need for enhanced LLM safety mechanisms.

In the business realm, Perplexity AI, an artificial intelligence search start-up, saw a seven-fold increase in monthly revenues and usage since the start of 2024, shifting its business model from subscriptions to advertising and bringing it into direct competition with Google. Meanwhile, OpenAI faced challenges as only two of its original 11 co-founders remained active at the company, raising questions about its direction, particularly regarding AI safety and research.

Mid-August saw Black Forest Labs release the Flux.1 family of text-to-image models, positioning itself as a significant player in the text-to-image model market. Researchers also developed TransAgents, a multi-agent workflow system for translating novels from Chinese to English, highlighting the potential of agentic workflows in complex tasks like literary translation.

The month concluded with Google introducing Gemini Live, an AI voice assistant planned to be embedded in its Android mobile operating system. This move was seen as a response to recent legal challenges and a strategic step in the battle for AI assistants' market share, which is likely to depend heavily on distribution.

What's happening in AI and regulation

The UK witnessed a surge in mass lawsuits over antitrust breaches, with notable cases including a £10.6 billion lawsuit against Google and a £2 billion claim against Amazon. This trend was attributed to the UK's adoption of an "opt-out" regime for class action lawsuits, with England accounting for nearly half of the class actions filed in Europe over the past five years.

Across the Channel, the EU took a strong stance against tech giants. Elon Musk's X (formerly Twitter) faced accusations of breaching the Digital Services Act regarding its blue "checkmark" practices, potentially facing fines of up to 6% of its total worldwide turnover.

The EU's privacy watchdog also requested Meta to voluntarily pause the training of future AI models on data in the region, leading to Meta's decision not to release future multimodal AI models in the European Union due to privacy law concerns.

In the United States, a federal judge dismissed key claims in a lawsuit against GitHub Copilot, Microsoft, and OpenAI, potentially setting a precedent for how AI developers can use training data. This ruling could provide clarity for AI developers on using training data and establish ethical use of code-completion tools trained on open-source code. Meta faced significant legal challenges, agreeing to pay £1.1 billion to the state of Texas to settle claims of harvesting biometric data without proper consent. This settlement, centred on Facebook's "tag suggestions" feature, marked the largest ever obtained from an action brought by a single US state.

A study by MIT's Data Provenance Initiative revealed that many websites are increasingly restricting access to content for AI training purposes, potentially limiting the availability of high-quality training data for AI systems. This trend could have significant implications for the future development and performance of AI models.

In a landmark antitrust case, US Federal Judge Amit Mehta ruled that "Google is a monopolist", highlighting three ways Google's dominance distorts competition. The UK Competition and Markets Authority also launched a formal merger inquiry into Amazon's £3.1 billion investment in AI start-up Anthropic, reflecting growing scrutiny of alliances between Big Tech and AI start-ups. Google faced further controversy when it was revealed that the company had made a secret deal with Meta to target advertisements for Instagram to 13- to 17-year-old YouTube users, disregarding Google's own rules prohibiting personalised ads for under-18s.

Apple, in an effort to comply with the EU's Digital Markets Act and avoid potential fines, announced further changes to its App Store rules in the EU. This marked the fourth time Apple had modified its EU business terms since initially moving to comply with the DMA earlier in the year.

Finally, a US federal judge ruled that Google violated antitrust law in a landmark case brought by the US Department of Justice. The ruling revealed that Google had spent billions annually on exclusivity deals, including a \$20 billion payment to Apple in 2022 alone to make Google's search engine the default on Safari.

What's happening in AI investment

Big tech companies like Google, Amazon, and Microsoft found themselves in a bit of a pickle. As they raced to develop more powerful AI systems, their efforts to be environmentally friendly took a hit. Google, for example, saw its carbon dioxide emissions jump by nearly half in just four years. But it's not all doom and gloom - these companies are working hard to find greener ways to power their AI dreams, and many believe AI itself could help fight climate change in the long run.

The AI boom has created a gold rush of sorts, with tech giants snapping up talented individuals and promising technologies from smaller AI companies. Amazon brought on board most of the team from a company called Adept AI, while Google did something similar with a chatbot company called Character.AI. Even Microsoft got in on the action, acquiring talent from a start-up called Inflection. This trend has some worried that only the biggest companies will end up benefiting from the AI revolution.

Money is pouring into AI like never before. The "Magnificent 7" tech companies - a group that includes giants like Amazon and Google - are spending eye-watering amounts on research and development. In fact, they're responsible for 40% of all the money spent on R&D by the top 500 companies in the US!

But it's not just tech companies getting rich off AI. Management consultancy firms are also cashing in, with companies like Boston Consulting Group expecting to make billions from AI-related work this year.

It's not all smooth sailing in the world of AI, however. The stock market had a bit of a wobble in July, with tech and AI stocks tumbling. Even Nvidia, a company that makes the special chips needed for AI and was briefly the world's most valuable company, saw its value drop by a whopping £586 billion.

There have been some growing pains in the industry. Nvidia and another company called TSMC ran into trouble making their next generation of super-powerful AI chips. This shows just how complicated and challenging it can be to push the boundaries of AI technology.

On a more positive note, the price of using some of the most advanced AI models has been falling. This is good news for developers and could lead to more innovation and wider use of AI technologies.

But developing cutting-edge AI isn't cheap. OpenAI, one of the leading companies in the field, is expected to lose billions of pounds this year despite making more money than ever before. This goes to show just how much investment is needed to stay at the forefront of AI technology.

What's happening in AI in Education and Employment

In July, the U.S. Department of Education released a helpful guide for people creating AI tools for schools. They called it "Designing for Education with Artificial Intelligence: An Essential Guide for Developers". This guide is meant to help teachers work with tech companies to make sure AI is used safely and effectively in schools. The main message? AI shouldn't be making decisions on its own - teachers should always be in charge. The guide also gives five important tips for making sure AI tools in schools are based on good teaching practices and don't treat some students unfairly.

Moving into August, we saw fascinating changes in how people look for jobs. More and more job seekers are using AI to help them apply for work. They're using clever AI tools to write their CVs (that's what we call resumes in the UK), prepare for interviews, and even get help during the interviews themselves! It's like having a super-smart assistant in your pocket.

This trend is partly because employers have been using AI to sort through job applications for a while now. So, job seekers are fighting fire with fire, you might say. It's creating a bit of a tech race in the job market, with both sides trying to outsmart each other with AI.

About half of all people looking for jobs are now using these AI tools. This has led to a big increase in the number of applications for each job - more than double what it used to be! This is making life quite tricky for the people doing the hiring, as they have many more applications to look through.

However, not everyone is happy about this AI trend in job hunting. Some big companies, like the major accountancy firms, have strict rules against using AI in job applications. They're saying it's not allowed at all.

South Korea's proposal to introduce AI-powered digital textbooks in schools has faced significant backlash from parents and academics. The government's plan, set to begin next year for pupils as young as eight, aims to shift away from rote learning. However, over 50,000 parents have signed a petition opposing the initiative, expressing concerns about children's overexposure to digital devices and potential negative impacts on brain development, concentration, and problem-solving abilities. Critics, including Professor Shin Kwang-Yeong of Chung-Ang University, argue that the government is implementing the plan too hastily without properly assessing potential side effects. Additional worries include the spread of misinformation, plagiarism risks, and the security of students' personal data. This opposition highlights a growing tension between technological innovation in education and concerns for children's overall wellbeing.

While all this is happening, the UK is facing a different kind of challenge in education. There's been a worrying drop in the number of older students (those aged 30 and over) applying for important jobs like teaching and nursing. Applications from these mature students fell by 15% compared to the

previous year. This is a big problem because the country needs more teachers and nurses. For example, last year, England fell short of its target for new teachers by 38%. The Royal College of Nursing says we need 11% more people applying for nursing courses every year until 2031 to have enough nurses for the future.

What's happening in AI ethics and societal impact

In July, some clever researchers made a worrying discovery about AI. They found that when AI systems are trained using computer-generated or 'synthetic' data, the results can be a bit bonkers. In one experiment, they started with information about medieval architecture (that's old castles and churches), but after the AI processed this information a few times, it somehow ended up talking about jackrabbits! This research shows that AI can sometimes go off track and make more and more mistakes over time, especially when it's not trained on real-world data.

Around the same time, Microsoft introduced a new technology called VASA-1. This clever bit of AI can create videos of people talking, even if those videos aren't real. While this might sound cool, it's also a bit scary. Some people are worried that this technology could be misused to create fake videos (called deepfakes) that could fool people. This has led to calls for better ways to spot these fake videos and for rules about how this kind of technology should be used.

Then, in August, something quite funny happened with ChatGPT. Some people in the UK who were asking questions in English suddenly started getting answers in Welsh! This even happened to people who don't speak a word of Welsh or have any connection to Wales. It turns out this mix-up was caused by a problem with how ChatGPT understands spoken languages. The company behind ChatGPT, OpenAI, admitted that their AI wasn't very good at Welsh because of some mistakes in how it was trained.

These events show us that while AI is becoming more and more clever, it still has some quirks and problems to work out. The synthetic data issue reminds us that AI needs good, real information to learn from. The VASA-1 technology highlights the need for careful thinking about how we use AI, especially when it comes to creating realistic-looking videos. And the Welsh language mix-up? Well, that just goes to show that even the smartest AI can sometimes get a bit confused!

What's happening in AI and Cybersecurity

In July, there was a bit of a tech disaster that affected millions of people. A company called CrowdStrike, which makes security software, accidentally sent out a faulty update. This update caused problems for a whopping 8.5 million Microsoft Windows devices! It was a bit like a domino effect, showing how interconnected our digital world is. This incident made people realise how risky it can be when we rely too much on just a few big tech companies. Did you know that just three companies - Google, Amazon, and Microsoft - control two-thirds of the cloud computing market? That's a lot of eggs in not many baskets!

This CrowdStrike mishap wasn't just a headache for computer users - it also caused a stir in the insurance world. Aon, a big insurance company, said this could be the biggest cyber insurance loss since a nasty computer virus called NotPetya caused havoc back in 2017. The damage could cost insurance companies anywhere from £780 million to several billion pounds. Ouch!

The drama continued into August when Delta Air Lines, a big American airline, threatened to sue CrowdStrike over the software update mess. Delta said the problem would cost them £390 million and hired some fancy lawyers to look into it. But CrowdStrike wasn't having any of it. They said it wasn't their fault that Delta's computers were down for so long and accused the airline of telling fibs about what happened.

While all this was going on in the business world, something quite different was happening in Ukraine. They've been developing a fleet of robot boats to help defend against Russian ships in the Black Sea. These aren't your average remote-controlled toys - these boats can navigate by themselves using AI. They're smaller and nimbler than big warships, which makes them hard to spot and stop. This clever use

of AI in warfare is getting attention from militaries around the world and could change how naval battles are fought in the future.

All of these events show us how AI and cybersecurity are becoming more and more important in our lives. The CrowdStrike incident reminds us that we need to be careful about relying too much on a few big tech companies and that we should always have backup plans for when things go wrong. The insurance industry's reaction shows how expensive these tech mishaps can be.

The dispute between Delta and CrowdStrike highlights how tricky it can be to figure out who's responsible when technology fails. And Ukraine's AI-powered boats show us that AI isn't just for chatbots and social media - it's also changing how countries defend themselves.

What's happening in the AI workforce and labour market

The chip manufacturing industry is facing a big problem: they're running out of workers! This isn't just a small hiccup; it's shaping up to be a major challenge for the future.

Let's break it down with some numbers. In the United States, the chip industry is expected to create more than 160,000 new job openings for engineers and technicians. That sounds like great news for job seekers, right? But here's the catch - only about 2,500 new workers are joining this industry each year. That's a huge gap!

If we do the maths, by the year 2029 (that's just a few years away), the US could be short of up to 146,000 workers in chip manufacturing. That's like trying to build a city but not having enough builders!

It's not just the US facing this problem. South Korea, another country famous for its technology, is also in a tight spot. They're looking at a shortage of 56,000 people in their chip industry by 2031.

Now, you might be wondering why this matters. Well, computer chips are incredibly important in our modern world. They're in almost everything electronic we use. If there aren't enough people to make these chips, it could slow down the production of all sorts of technology. This could affect everything from the price of our gadgets to how quickly new technologies can be developed and released.

This situation shows us a few important things:

1. The technology industry is growing really fast - maybe faster than our education systems can keep up with.
2. There's a big opportunity for young people (and even older folks looking for a career change) to get into this field. If you're interested in technology and engineering, the chip industry could be a great place to look for a job in the future.
3. Countries and companies might need to think of creative ways to train more people quickly for these jobs. This could mean new education programmes, apprenticeships, or even using technology itself to help train people faster.

Sources

AI Development and Industry

July 2024

- 11 July: Noah Yuval Harari warns of AI's potential to destroy trust between people and shift trust to AI systems, particularly in finance and politics.
 - Harari suggests AI could make the financial system too complex for human comprehension.
 - He warns of potential political implications, with AI potentially controlling important aspects of the world and politics.
 - Harari emphasises the danger of giving control to non-human intelligence that cannot be regulated or supervised. [Source: Instagram Reel,

11/07/2024] <https://www.instagram.com/reel/C9h74u4Nkxz/?igsh=MThhcWJyNm5wdnluYQ==>

- 11 July: The US Department of Education releases guidance for AI ed-tech developers, emphasising responsible development practices.
 - The guidance aims to help educators work with vendors and tech developers to manage risks associated with AI-driven innovations in schools.
 - It emphasises that AI should not make decisions unchecked by educators and that AI tools should be based on evidence-based practices.
 - The document provides five key recommendations for vendors and educators, including designing with teaching and learning in mind and taking steps to remove or mitigate bias. [Source: GovTech, 11/07/2024] <https://www.govtech.com/education/k-12/u-s-ed-issues-guidance-for-ai-ed-tech-developers>
- 11 July: Anthropic introduces 'Artifacts', a new feature for Claude 3.5 Sonnet that enhances the user interface for working with large language models.
 - Displays outputs in a separate window of the web interface.
 - Artifacts can include documents, code snippets, HTML pages, vector graphics, or JavaScript visualisations.
 - Users can enable artifacts from the 'feature preview' dropdown in their profile menu at Claude.ai.
 - Artifacts are typically at least 15 lines long and can be viewed as code or rendered display.
 - Makes working with large language models more fluidly interactive, especially for tasks like coding and image generation. [Source: The Batch, Issue 257] <https://www.deeplearning.ai/the-batch/issue-257/>
- 11 July: Researchers develop Gradient Low-Rank Projection (Galore), a new method for memory-efficient AI model training.
 - Works well for both fine-tuning and pretraining, unlike Low-Rank Adaptation (LoRA).
 - Used to pretrain a 7B parameter transformer using a consumer-grade Nvidia RTX 4090 GPU.
 - Achieved similar performance to traditional methods while using significantly less memory.
 - Could make AI model training more accessible, potentially allowing for training of large models on consumer-grade hardware. [Source: The Batch, Issue 257] <https://www.deeplearning.ai/the-batch/issue-257/>
- 18 July: Researchers at Radboud University evaluate AI model openness, providing guidelines for developers and policymakers.
 - Evaluated AI model openness based on 14 characteristics across availability, documentation, and access.
 - Openness is crucial for innovation in AI, enabling developers to build on each other's work.
 - The study provides clear guidelines for developers and policymakers on what constitutes an 'open' AI model.
 - Has growing regulatory implications, such as in the European Union's AI Act. [Source: The Batch, Issue 258] <https://www.deeplearning.ai/the-batch/issue-258/>
- 18 July: Artificial Analysis introduces the Text to Image Arena leaderboard, ranking text-to-image models.
 - Ranks text-to-image models based on head-to-head matchups judged by the public.
 - Midjourney v6 currently leads in quality but lags in speed.
 - Aggregating user preferences measures quality, while personalised leaderboards cater to individual tastes.
 - Could potentially feed into algorithms like reinforcement learning from human feedback to improve image generators. [Source: The Batch, Issue 258] <https://www.deeplearning.ai/the-batch/issue-258/>
- 18 July: University of Oxford researchers propose a method to identify LLM hallucinations.
 - Calculates entropy based on the distribution of generated meanings rather than sequences of words.

- Crucial for building trust in AI systems and increasing adoption, particularly in fields like medicine or law.
- Helps researchers identify common circumstances where hallucinations occur, potentially leading to improvements in future models. [Source: The Batch, Issue 258] <https://www.deeplearning.ai/the-batch/issue-258/>
- 24 July: JPMorgan Chase rolls out LLM Suite, a generative AI product, to employees in its asset and wealth management division.
 - The bank is pitching LLM Suite as capable of doing the work of a research analyst, helping with writing, idea generation, and summarising documents.
 - About 50,000 employees (15% of staff) now have access to LLM Suite.
 - JPMorgan developed the proprietary platform in-house due to data security regulations. [Source: Reuters, 26/07/2024] <https://www.reuters.com/technology/artificial-intelligence/jpmorgan-launches-in-house-chatbot-ai-based-research-analyst-ft-reports-2024-07-26/>
- 24 July: AI advancements significantly accelerate capabilities in the £58bn robotics sector.
 - Improved computer vision and spatial reasoning have allowed robots to gain greater autonomy in various environments.
 - Deep learning models have enabled robots to be more adaptive and reactive to unexpected physical challenges.
 - The global market for humanoid robots is valued at over £1bn and growing 20% annually. [Source: Financial Times, 24/07/2024] <https://on.ft.com/4dgQRMB>
- 25 July: OpenAI launches SearchGPT, an experimental online search tool, challenging Google's dominance in the search market.
 - SearchGPT offers a rail of clickable links to external websites, similar to traditional search engines, in addition to generating answers.
 - The product was developed with feedback from publishers that OpenAI has recently signed deals with.
 - OpenAI faces challenges with AI-generated search, including the tendency for chatbots to provide inaccurate information or "hallucinate" facts. [Source: Computerworld, 26/07/2024] <https://www.computerworld.com/article/3477894/openais-ai-powered-searchgpt-is-set-to-challenge-googles-web-search-dominance.html>
- 25 July: Lord Patrick Vallance, the UK's new science minister, states that cyber security must be a national priority to safeguard important data troves.
 - The government has announced a cyber security and resilience bill in response to recent high-profile ransom attacks and IT breakdowns.
 - Amazon Web Services will provide cloud computing storage access worth about £8 million to the UK Biobank genetic database.
 - Vallance emphasises the need for a balanced approach to cyber security that maintains access for groundbreaking research. [Source: Financial Times, 25/07/2024] <https://on.ft.com/4c43jy1>
- 25 July: OpenAI releases GPT-4o mini, a smaller, cost-effective version of their multimodal GPT-4o model.
 - More cost-effective than its larger counterpart and outperforms similar-sized models from competitors.
 - Accepts text and image inputs, with video and audio capabilities coming soon.
 - Has a context window of 128,000 tokens.
 - Significantly cheaper to use via API than the full GPT-4o.
 - Sets a new standard for smaller, more efficient AI models. [Source: The Batch, Issue 259] <https://www.deeplearning.ai/the-batch/issue-259/>
- 25 July: Microsoft researchers develop VASA-1, an advanced system for generating expressive talking-head videos.
 - Uses a single portrait image and an audio clip.
 - Animates the entire head as a whole, resulting in more natural and expressive outputs.
 - Uses separate embeddings for facial structure, expression, and head position.
 - Maintains consistent facial features while generating appropriate motions.

- Represents a significant advancement in synthetic media generation. [Source: The Batch, Issue 259] <https://www.deeplearning.ai/the-batch/issue-259/>

August 2024

- 1 August: Meta releases Llama 3.1, a family of open-source language models rivalling top proprietary models in performance.
 - Used extensive data curation and generation techniques to improve capabilities.
 - Demonstrates the critical importance of data-centric AI in developing high-performing language models.
 - Provides valuable insights into systematically engineering training data.
 - Achieved better performance than proprietary models on multiple benchmarks.
 - Highlights the growing potential of open-weights models. [Source: The Batch, Issue 260] <https://www.deeplearning.ai/the-batch/issue-260/>
- 1 August: OpenAI tests SearchGPT, an AI-powered search engine combining web crawling with conversational AI capabilities.
 - Aims to compete with Google and Microsoft Bing.
 - Offers direct answers to queries and a conversational interface for follow-up questions.
 - Includes licensing content from trusted sources.
 - Represents a significant step in enhancing web search with AI capabilities. [Source: The Batch, Issue 260] <https://www.deeplearning.ai/the-batch/issue-260/>
- 1 August: Microsoft researchers introduce AgentInstruct, a framework for producing synthetic data to fine-tune large language models.
 - Uses agentic workflows to generate diverse, high-quality training data.
 - Offers a pattern for AI engineers to build synthetic datasets for various tasks.
 - Potentially improves model performance across a range of applications.
 - Success demonstrated by the performance of the resulting Orca-3 model.
 - Highlights the value of methodical, diverse data generation in AI development. [Source: The Batch, Issue 260] <https://www.deeplearning.ai/the-batch/issue-260/>
- 8 August: Researchers discover that ASCII art can bypass safety measures in large language models.
 - Technique called ArtPrompt successfully circumvented guardrails in several major LLMs.
 - Adds to the growing body of literature on LLM jailbreak techniques.
 - Highlights the lack of robust mechanisms to prevent various jailbreak methods.
 - Underscores the need for additional input and output screening systems to enhance LLM safety. [Source: The Batch, Issue 261] <https://www.deeplearning.ai/the-batch/issue-261/>
- 9 August: Perplexity AI, an artificial intelligence search start-up, sees a seven-fold increase in monthly revenues and usage since the start of 2024.
 - The company answered approximately 250 million questions in the last month, compared to 500 million for the whole of 2023.
 - Perplexity has raised £195 million in a new funding round, tripling its valuation to £2.3 billion.
 - The start-up is shifting its business model from subscriptions to advertising, bringing it into direct competition with Google. [Source: Verdict, 12/08/2024] <https://www.verdict.co.uk/ai-search-start-up-perplexity-challenges-googles-search-dominance/?cf-view>
- 10 August: Only two of OpenAI's original 11 co-founders remain active at the company.
 - The exodus of founders followed the November 2023 attempted boardroom coup against CEO Sam Altman.
 - Some founders have joined competitors, while others have started their own AI companies.
 - The departures have raised questions about OpenAI's direction, particularly regarding AI safety and research. [Source: Observer, 12/08/2024] <https://observer.com/2024/07/openai-founders-career/>
- 15 August: Black Forest Labs releases the Flux.1 family of text-to-image models.
 - Includes both proprietary and open-source versions.

- Top models outperform competitors in internal tests.
 - Covers all tiers of text-to-image models: large commercial models, open-source offerings, and smaller models for local use.
 - Positions Flux.1 as a significant player in the text-to-image model market. [Source: The Batch, Issue 262] <https://www.deeplearning.ai/the-batch/issue-262/>
- 15 August: Researchers develop TransAgents, a multi-agent workflow system for translating novels from Chinese to English.
 - Uses multiple LLM instances to mimic different roles in a translation company.
 - Literary translation remains a challenging frontier for machine translation.
 - Agentic workflow approach breaks down the task into subtasks for different LLM instances.
 - Results appear to appeal to human judges.
 - Highlights the need for new ways to measure the quality of literary translations.
 - Raises important research questions about optimal task division in agentic workflows. [Source: The Batch, Issue 262] <https://www.deeplearning.ai/the-batch/issue-262/>
- 16 August: Google introduces Gemini Live, an AI voice assistant that could become a primary way for users to interact with smartphones.
 - Google plans to embed Gemini in its Android mobile operating system, potentially reaching about 70% of smartphone users.
 - This move is seen as a response to the recent US court ruling declaring Google an illegal monopoly in internet search.
 - The battle for AI assistants' market share will likely depend heavily on distribution, similar to how Google dominated search. [Source: Tech Crunch, 13/08/2024] <https://techcrunch.com/2024/08/13/gemini-live-googles-answer-to-chatgpts-advanced-voice-mode-launches/>

AI Regulation and Legal Issues

July 2024

- 11 July: Mass lawsuits over antitrust breaches are becoming more common in the UK.
 - A £10.6bn lawsuit against Google for alleged anti-competitive behaviour given the green light.
 - A £2bn claim against Amazon for allegedly exploiting sellers on its platform is being pursued.
 - The UK has adopted an "opt-out" regime for class action lawsuits.
 - England accounts for nearly half of the class actions filed in Europe over the past five years. [Source: Tech Policy Press, 04/06/2024] <https://www.techpolicy.press/a-new-british-law-charts-a-course-towards-reining-in-big-tech/>
- 12 July: The EU accuses Elon Musk's X (formerly Twitter) of breaching the Digital Services Act regarding its blue "checkmark" practices.
 - The EU's preliminary findings suggest X's practice of allowing anyone to pay for a blue checkmark is deceptive.
 - If confirmed, X could face fines of up to 6% of its total worldwide turnover.
 - The EU also claims X has not complied with DSA requirements in areas such as transparency in advertising and data access for researchers. [Source: Scottish Legal News, 15/07/2024] <https://www.scottishlegal.com/articles/eu-takes-action-against-blue-checkmarks-for-all-on-elon-musks-x-platform>
- 18 July: A U.S. federal judge dismisses key claims in a lawsuit against GitHub Copilot.
 - Dismissed key claims of copyright infringement and unfair profit in a class-action lawsuit against GitHub Copilot, Microsoft, and OpenAI.
 - Ruled that plaintiffs failed to provide concrete evidence that Copilot could generate substantial copies of code.
 - Could have broader implications for how AI developers can use training data.
 - Supports the view that AI developers may have a broad right to use data for training models, even if protected by copyright.
 - Could provide clarity for AI developers on using training data and establish ethical use of code-completion tools trained on open-source code. [Source: The Batch, Issue 258] <https://www.deeplearning.ai/the-batch/issue-258/>

- 23 July: Meta warns that the EU's approach to regulating AI may prevent the bloc from accessing cutting-edge services.
 - The EU's privacy watchdog has requested Meta to voluntarily pause the training of future AI models on data in the region.
 - Meta will not ship multimodal models in the EU, and future AI releases may have limited availability in Europe.
 - The company cites uncertainty over whether training AI models on consumer data complies with EU's General Data Protection Regulation (GDPR). [Source: Irish Times, 24/06/2024] <https://www.irishtimes.com/business/2024/06/24/meta-backs-new-round-of-ai-supports/>
- 25 July: Meta decides not to release future multimodal AI models in the European Union due to privacy law concerns.
 - Follows objections from EU member states regarding Meta's data collection practices.
 - The company will continue to offer text-only models, including Llama 3.1, in the EU.
 - Part of a broader trend of tech companies grappling with EU regulations.
 - Highlights growing tension between rapid AI development and regulatory frameworks, particularly in the EU.
 - Could potentially slow down AI innovation and adoption in Europe, creating a technological gap between regions with different regulatory approaches. [Source: The Batch, Issue 259] <https://www.deeplearning.ai/the-batch/issue-259/>
- 30 July: Meta agrees to pay £1.1 billion to the state of Texas to settle claims of harvesting biometric data without proper consent.
 - This is the largest settlement ever obtained from an action brought by a single US state.
 - The lawsuit centred on Facebook's "tag suggestions" feature, launched in 2011, which scanned facial geometry to recommend tags in photos and videos.
 - Meta shut down its facial recognition system in 2021, citing legal "uncertainty". [Source: Reuters, 31/07/2024] <https://www.reuters.com/technology/cybersecurity/meta-platforms-pay-1.4-bln-settle-texas-lawsuit-over-facial-recognition-data-2024-07-30/>

August 2024

- 1 August: A study finds many websites are restricting access to content for AI training purposes.
 - Study by MIT's Data Provenance Initiative found many websites increasingly restricting access to content for AI training purposes.
 - Over the past year, websites responsible for half of all tokens in popular training datasets changed their terms of service or robots.txt files.
 - Changes forbid crawlers or AI training use.
 - Could limit the availability of high-quality training data for AI systems.
 - May affect both commercial AI developers and academic research.
 - Could have significant implications for future development and performance of AI models. [Source: The Batch, Issue 260] <https://www.deeplearning.ai/the-batch/issue-260/>
- 8 August: US Federal Judge Amit Mehta rules that "Google is a monopolist" in an antitrust case brought by the US Department of Justice.
 - The ruling came after nearly four years of court proceedings, involving millions of pages of submissions, 3,500 exhibits, and dozens of witness testimonies.
 - Judge Mehta highlighted three ways Google's dominance distorts competition: its grip on the search market, its business model compromising user privacy, and its payments to companies like Apple for default search distribution.
 - Possible remedies could include fines, scrapping Google's distribution deals, or more radical options like breaking up Alphabet. [Source: EPRA, 14/08/2024] https://www.epra.org/news_items/usa-an-illegal-monopoly-for-google-according-to-a-federal-judge
- 8 August: The UK Competition and Markets Authority launches a formal merger inquiry into Amazon's £3.1 billion investment in AI start-up Anthropic.

- The CMA will decide by 4 October whether to escalate the inquiry into a more in-depth "Phase 2" investigation.
- Unlike Microsoft's deal with OpenAI, Amazon is not Anthropic's exclusive cloud provider and doesn't hold a board seat.
- Regulators worldwide are increasingly scrutinising alliances between Big Tech and AI start-ups. [Source: Independent, 08/08/2024] <https://www.independent.co.uk/business/competition-watchdog-launches-inquiry-into-amazon-s-partnership-with-anthropic-b2593347.html>
- 8 August: Google and Meta made a secret deal to target advertisements for Instagram to 13- to 17-year-old YouTube users, disregarding Google's own rules prohibiting personalised ads for under-18s.
 - The campaign deliberately targeted a group labelled as "unknown" in Google's advertising system, which was known to skew towards under-18s.
 - The companies worked with Spark Foundry, a subsidiary of Publicis, to launch pilot programmes in Canada and the US.
 - Google initiated an investigation after being contacted by the Financial Times and has since cancelled the project. [Source: Ars Technica, 08/08/2024] <https://arstechnica.com/tech-policy/2024/08/google-and-meta-ignored-their-own-rules-in-secret-teen-targeting-ad-deals/>
- 9 August: Apple announces further changes to its App Store rules in the EU, aiming to comply with the Digital Markets Act and avoid potential fines.
 - This is the fourth time Apple has modified its EU business terms since initially moving to comply with the DMA earlier in the year.
 - The changes include a new fee structure and eased rules around how developers can display links within their apps.
 - If found non-compliant with the DMA, companies like Apple could face fines of up to 10% of their global turnover. [Source: BBC, 24/06/2024] <https://www.bbc.co.uk/news/articles/c5111qxl2nro>
- 11 August: A US federal judge ruled that Google violated antitrust law in a landmark case brought by the US Department of Justice.
 - Google was found to have spent billions annually on exclusivity deals with wireless carriers, browser developers, and device manufacturers.
 - Apple received \$20 billion from Google in 2022 alone to make Google's search engine the default on Safari.
 - Google controls 90% of the search market, rising to 95% for mobile devices. [Source: Pro Market, 20/08/2024] <https://www.promarket.org/2024/08/20/the-us-google-search-case-is-really-about-monopolizing-the-future/>

AI Market and Investment

July 2024

- 11 July: The AI boom is challenging big tech companies' efforts to reach greenhouse gas emissions targets.
 - Google's total carbon dioxide emissions rose nearly 50% between 2019 and 2023 to 14.3 million tonnes.
 - Three-quarters of total emissions are associated with purchases including data-centre hardware and construction.
 - Google is working to reduce emissions through various means, including purchasing low-emissions electricity and developing more efficient hardware.
 - Similar trends seen with Amazon and Microsoft, whose emissions have also increased significantly.
 - Despite challenges, AI is seen as a potential tool to mitigate climate change in various sectors. [Source: The Batch, Issue 257] <https://www.deeplearning.ai/the-batch/issue-257/>
- 11 July: Amazon acquires most of Adept AI's staff and technology.
 - Amazon hired most of the staff from agentic-AI specialist Adept AI.
 - Licensed Adept's models, datasets, and other technology non-exclusively.

- Two-thirds of Adept's former employees, including CEO David Luan and four co-founders, joined Amazon's artificial general intelligence (AGI) autonomy team.
 - The team will focus on building agents that can automate software workflows.
 - Part of Amazon's broader strategy to compete in AI for both businesses and consumers. [Source: The Batch, Issue 257] <https://www.deeplearning.ai/the-batch/issue-257/>
- 24 July: Alphabet's revenues increased by 14% to £66.2 billion in the second quarter, beating analysts' estimates.
 - Advertising revenue, which accounts for most of Google's income, grew 11% to £50.5 billion.
 - Google Cloud's services business saw a 29% increase to £8.1 billion.
 - Capital expenditure increased to £10.2 billion, nearly double the amount spent in the same period in 2023. [Source: Reuters, 24/07/2024] <https://www.reuters.com/technology/google-parent-alphabet-beats-quarterly-revenue-estimates-2024-07-23/>
- 24 July: US stock indices experienced their worst day in over 18 months.
 - The S&P 500 fell 2.3% and the Nasdaq Composite dropped 3.6%.
 - The sell-off was driven by disappointing results from Tesla and Alphabet, as well as a broader decline in tech and AI stocks.
 - Tesla's stock fell 12.3%, its worst daily performance since 2020, after reporting profits well below expectations.
 - Other tech giants, including Nvidia, Microsoft, Apple, and Meta, also experienced significant declines. [Source: Guardian, 05/08/2024] <https://www.theguardian.com/business/article/2024/aug/05/us-stock-market-recession-dow-close>
- 25 July: Venture capital firms, particularly Andreessen Horowitz, are stockpiling high-end GPUs to attract AI startups.
 - AI6Z plans to acquire over 20,000 GPUs, including top-tier Nvidia H100s.
 - Offer these resources to portfolio companies at reduced rates or in exchange for equity.
 - Other investors and companies, such as Ex-GitHub CEO Nat Friedman and Microsoft, are also providing GPU access to startups they support.
 - Demonstrates intense competition among venture capital firms to invest in promising AI startups.
 - May influence the direction of AI development, as startups might tailor their projects to leverage available GPU power. [Source: The Batch, Issue 259] <https://www.deeplearning.ai/the-batch/issue-259/>
- 30 July: Management consultancy firms benefit significantly from the AI revolution.
 - Boston Consulting Group estimates that 20% of its revenues this year will come from AI-related work, potentially amounting to £2 billion.
 - Accenture has booked £1.6 billion of AI-related projects year to date.
 - Goldman Sachs estimates £780 billion of capital expenditure will be invested in AI infrastructure in the coming years. [Source: Financial Times, 30/07/2024] <https://on.ft.com/3SyXr9d>
- 31 July: The "Magnificent 7" tech companies account for 18% of the S&P 500's total capital expenditures in the last fiscal year, up from 5% ten years ago.
 - The Mag 7's share of research and development spending was 40% of the S&P's total, or £189 billion last year.
 - Combined capex and R&D for the Mag 7 totalled £328 billion last year.
 - Amazon leads in investment, with £41 billion in capex and £67 billion in R&D last year. [Source: Financial Times, 31/07/2024] <https://on.ft.com/3YnlyLJ>
- 31 July: Nvidia's shares close down 7%, with the company's market capitalisation falling by almost £586bn since briefly becoming the world's most valuable publicly traded company.
 - Other chip stocks, including Arm, also experienced significant drops.
 - Despite recent declines, Nvidia and Arm are still more than double their value compared to a year ago.

- Traders are concerned that profit expectations for AI-involved companies are too high and that capital spending is outpacing returns. [Source: Financial Times, 31/07/2024] <https://on.ft.com/4d0k6DK>

August 2024

- 5 August: Nvidia and TSMC face production challenges with the next generation of Nvidia's most powerful AI chips, the Blackwell family.
 - Nvidia's shares fell as much as 15% in early trading, while TSMC closed down 10% on the Taiwan stock exchange.
 - The production difficulties are related to the interposer, a layer connecting different dies in the complex chips.
 - Analysts project £780 billion of spending on AI data centres over the next five years. [Source: Reuters, 05/08/2024] <https://www.reuters.com/technology/artificial-intelligence/nvidias-ai-chip-demand-face-limited-impact-potential-production-delay-analysts-2024-08-05/>
- 8 August: Google acquires Character.AI co-founders and technology.
 - Google hired the co-founders and several employees of Character.AI, a chatbot company.
 - Acquired nonexclusive rights to its technology.
 - Character.AI will shift focus to using open-source models for chatbot development.
 - Highlights significant costs associated with developing cutting-edge foundation models.
 - Demonstrates trend of essential team members from startups moving to AI giants.
 - Startups need to adapt business strategies in a changing market.
 - Established companies benefit from acquiring entrepreneurial talent. [Source: The Batch, Issue 261] <https://www.deeplearning.ai/the-batch/issue-261/>
- 14 August: Major tech companies acquire talent from three promising AI start-ups in the past six months.
 - Inflection, Character.AI, and Adept were acquired by Microsoft, Google, and Amazon respectively.
 - These start-ups had collectively raised over £1.6 billion in funding before being acquired.
 - The deals typically involve hiring founders and key staff, as well as licensing products, rather than full acquisitions.
 - Venture capitalists are concerned that the main beneficiaries of the AI boom will be large tech companies that can afford the multibillion-dollar costs of developing cutting-edge AI systems. [Source: AFP/Yahoo, 28/07/2024] <https://sg.news.yahoo.com/ai-startups-swap-independence-big-034335119.html>
- 15 August: Various large language models see significant price reductions.
 - Includes GPT-4o, Llama 3.1 405B, Gemini 1.5 Flash, and DeepSeek v2.
 - Models offer a wide range of price and performance options.
 - Creating an extraordinary range of choices for developers.
 - Putting pressure on makers of foundation models.
 - Models that can't match the best large models in performance or the best small models in cost are facing challenges.
 - Likely to accelerate innovation and adoption of LLM technologies. [Source: The Batch, Issue 262] <https://www.deeplearning.ai/the-batch/issue-262/>
- 15 August: OpenAI faces potential significant losses in 2024 despite rapid revenue growth.
 - Reports suggest OpenAI's operating expenses could reach \$8.5 billion in 2024.
 - Revenue expected to be between \$3.5 billion and \$4.5 billion.
 - Could potentially result in a loss of \$4-5 billion.
 - Major costs include processing power, model training, and personnel.
 - Despite rapid growth in revenue, OpenAI faces substantial development, training, and operational costs.
 - Financial outlook may improve due to rising revenue, improving cost-effectiveness of models, and falling inference costs.

- Highlights significant investment required to develop cutting-edge AI technologies. [Source: The Batch, Issue 262] <https://www.deeplearning.ai/the-batch/issue-262/>

AI in Education and Employment

July 2024

- 11 July: The U.S. Department of Education released guidance titled "Designing for Education with Artificial Intelligence: An Essential Guide for Developers".
 - The guidance aims to help educators work with vendors and tech developers to manage risks associated with AI-driven innovations in schools.
 - It emphasises that AI should not make decisions unchecked by educators and that AI tools should be based on evidence-based practices.
 - Five key recommendations are provided for vendors and educators, including designing with teaching and learning in mind and taking steps to remove or mitigate bias. [Source: GovTech, 11/07/2024] <https://www.govtech.com/education/k-12/u-s-ed-issues-guidance-for-ai-ed-tech-developers>

August 2024

- 8 August: Job seekers increasingly use AI tools for applications and interviews.
 - Tools used for resume writing, interview preparation, and real-time assistance during interviews.
 - Counters employers' use of AI in screening and hiring processes.
 - Creating an 'automation arms race' in the hiring process.
 - Making the hiring process more efficient for both parties.
 - Need for AI-powered solutions that can better match employers and candidates.
 - Such solutions could potentially benefit both sides of the hiring process. [Source: The Batch, Issue 261] <https://www.deeplearning.ai/the-batch/issue-261/>
- 13 August: Approximately half of all job seekers are using artificial intelligence tools to apply for roles.
 - Candidates are increasingly using generative AI to write CVs, cover letters, and complete assessments.
 - The number of candidates per job has more than doubled, making it harder for recruiters to sift through applications.
 - Many large employers, including the Big Four accountancy firms, have a zero-tolerance policy towards the use of AI in applications. [Source: Readwrite, 14/08/2024] <https://readwrite.com/recruiters-ai-job-applications-cvs/>
- 14 August: The UK experiences a concerning decline in applications from mature students for strategically important professions.
 - Applications from students aged 30 and over dropped by 15% in 2024 compared to the previous year, with just under 34,500 applications made.
 - In 2023-24, postgraduate teacher recruitment in England was 38% below target.
 - The Royal College of Nursing stated that nursing course applications need to rise by an average of 11% annually until 2031 to meet NHS England's long-term workforce plan. [Source: Research Professional News, 16/08/2024] <https://www.researchprofessionalnews.com/rr-news-uk-universities-2024-8-results-day-2024-mature-student-numbers-continue-to-plummet/>
- 18 August: South Korea's plan to introduce AI-powered digital textbooks in schools faces significant opposition from parents and academics.
 - The initiative, set to begin next year for pupils as young as eight, aims to shift away from traditional rote learning.
 - Over 50,000 parents have signed a petition against the plan, citing concerns about children's overexposure to digital devices.
 - 54% of state schoolteachers surveyed by the Korean Federation of Teachers' Associations expressed support for the initiative.
 - Critics argue the government is implementing the plan too hastily without properly assessing potential side effects, including risks of misinformation and data security concerns. [Source: Tech Crunch,

18/08/2024] <https://techcrunch.com/2024/08/18/south-koreas-ai-textbook-program-faces-skepticism-from-parents/>

AI Ethics and Societal Impact

July 2024

- 24 July: New research suggests that using computer-generated 'synthetic' data to train AI models could lead to nonsensical results.
 - The study, published in Nature, shows that using synthetic data could cause rapid degradation of AI models.
 - One trial using synthetic input text about medieval architecture devolved into a discussion about jackrabbits after fewer than 10 generations.
 - The research highlights the tendency of AI models to 'collapse' over time due to accumulation and amplification of mistakes. [Source: Live Science, 09/08/2024] <https://www.livescience.com/technology/artificial-intelligence/ai-models-trained-on-ai-generated-data-could-spiral-into-unintelligible-nonsense-scientists-warn>
- 25 July: Microsoft's VASA-1 technology for generating expressive talking-head videos raises ethical concerns.
 - Raises ethical concerns about potential misuse for creating deepfakes or misleading content.
 - Highlights need for robust detection methods and ethical guidelines. [Source: The Batch, Issue 259] <https://www.deeplearning.ai/the-batch/issue-259/>

August 2024

- 15 August: Some UK users of ChatGPT experience a glitch where the AI responds in Welsh to English-language queries.
 - This issue has occurred even for users who don't understand Welsh or have any connection to Wales.
 - OpenAI admitted in a research paper that ChatGPT's performance in Welsh was "much worse than expected" due to misidentified training data.
 - The company attributes the problem to limitations in ChatGPT's voice transcription system, Whisper, which sometimes confuses languages. [Source: Yahoo/Wales Online, 17/08/2024] <https://uk.news.yahoo.com/openais-chatgpt-starts-speaking-welsh-142608441.html>

AI and Cybersecurity

July 2024

- 23 July: A faulty software update from CrowdStrike affected 8.5 million Microsoft Windows devices, causing widespread disruptions.
 - The incident highlighted the vulnerability of interconnected digital systems and the risks of concentrated tech providers.
 - Three companies (Google, Amazon, and Microsoft) account for two-thirds of the cloud provider market.
 - The article emphasises the need for businesses and governments to build resilience against IT failures. [Source: Forbes, 19/07/2024] <https://www.forbes.com/sites/edwardsegal/2024/07/19/global-it-outage-provides-several-crisis-management-lessons/>
- 23 July: The insurance industry is preparing for potential losses due to the CrowdStrike incident.
 - Aon, a major insurance broker, considers this the most significant cyber insurance loss event since the NotPetya malware attacks of 2017.
 - Estimates for insured losses range from upwards of £780 million to potentially several billion pounds.
 - The incident highlights the interconnected nature of software ecosystems and the systemic nature of cyber risk. [Source: Insurance Business Magazine,

09/08/2024] <https://www.insurancebusinessmag.com/us/news/breaking-news/crowdstrike-why-did-insurers-get-off-quite-lightly-500747.aspx>

August 2024

- 5 August: CrowdStrike responds to Delta Air Lines' threat of litigation over the botched software update.
 - The cyber security firm denies responsibility for Delta's IT decisions and prolonged disruption.
 - Delta claims the disruption will cost it £390 million and has hired litigation firm Boies Schiller Flexner.
 - CrowdStrike argues that Delta created a "misleading narrative" about the company being "grossly negligent". [Source: BBC, 09/08/2024] <https://www.bbc.co.uk/news/articles/c6284e7r7d7o>
- 8 August: Ukraine develops a fleet of autonomous robotic watercraft to counter Russian naval forces.
 - Ukraine developed a fleet of robotic watercraft, including surface and underwater vessels.
 - Used to counter Russian naval forces in the Black Sea.
 - Drones have autonomous navigation capabilities.
 - Significantly impacted naval balance in the region.
 - Success of smaller, nimbler drones in warfare attracting attention from major military powers.
 - Offer different tactical and strategic opportunities compared to larger autonomous vessels.
 - Could potentially change future naval warfare strategies. [Source: The Batch, Issue 261] <https://www.deeplearning.ai/the-batch/issue-261/>

AI Workforce and Labour Market

August 2024

- 13 August: The chip manufacturing industry faces a looming shortage of engineers and technicians.
 - The US chip industry is expected to create over 160,000 new job openings in engineering and technician support, but only about 2,500 new workers join the industry annually.
 - By 2029, the US could face a shortage of up to 146,000 workers in the chip manufacturing sector.
 - South Korea is projected to have a labour shortage of 56,000 people in the chip industry by 2031. [Source: Observer, 12/08/2024] <https://observer.com/2024/08/ai-semiconductor-labor-shortage-us-mckinsey/>
- 18 August: A study finds that most Fortune 500 companies view AI as a 'risk factor' to their businesses.
 - 56% of Fortune 500 companies cited AI as a "risk factor" in their most recent annual reports, up from just 9% in 2022.
 - Only 33 out of 108 companies that specifically discussed generative AI saw it as an opportunity.
 - Over 90% of the largest US media and entertainment companies, and 86% of software and technology groups, identified AI systems as a business risk this year.
 - Potential risks cited include increased competition, reputational issues, operational challenges, and concerns about AI's impact on human rights, employment, and privacy. [Source: Fortune, 18/08/2024] <https://fortune.com/2024/08/18/ai-risks-fortune-500-companies-generative-artificial-intelligence-annual-reports/>

Items of Interest

- [OpenAI working on new reasoning technology under code name 'Strawberry'](#)

- [Private School in Austin, TX and Miami, FL: Alpha School](#)
- [Introducing the Coalition for Secure AI, an OASIS Open Project](#)
- [Academic authors 'shocked' after Taylor & Francis sells access to their research to Microsoft AI](#)
- [Videogame voice actors are officially on strike over AI: 'We refuse this paradigm'](#)
- [On Speaking to AI](#)
- [AI will never substitute teachers. This failed experiment shows why](#)
- [White House Targets Time-Wasting Chatbots in 'Time is Money' Campaign](#)
- [AI potential 'squandered by universities' risk-focused approach](#)