AI NEWS VOL 8

The US-Led AI Manhattan Project

By Captain Yar / July 4, 2025

I've been tracking the AI race between companies, but a new analysis from *Epoch AI* just put things on a completely different level.

They're talking about a full-blown, US-led *AI Manhattan Project*, and it's a total game-changer. The idea is to mobilize national resources, just like the Apollo program, to massively accelerate AI progress.

The scale is just mind-blowing. Here's the breakdown:

- **The Goal:** Train a model that's a staggering 10,000x larger than GPT-4 by late 2027.
- **The Resources:** It would require an Apollo-level investment to fund an estimated 27 million GPUs. Insane, right?
- **The Power:** This initiative would need so much energy that it might require using the Defense Production Act to fast-track the construction of power plants.

This isn't just a wild idea, either. A U.S. commission has already recommended it as a top priority for achieving AGI.

This completely reframes the race for AGI. It's not just a corporate marathon anymore; it's a national sprint with humanity-altering stakes.

JUNE 12, 2025

A foundation for physical AI: Battery-free RFID sensing system offers real-time, reliable data

by University of California - San Diego edited by Sadie Harley, reviewed by Robert Egan

Using a smartphone-based augmented reality app, conference attendees at IEEE RFID 2025 see in real time how UC San Diego researchers can harness RFID tags to show the amount of force applied or visualize the moisture content in a pot of soil. Credit: Nagarjun Bhat, UC San Diego
What if the same RFID "smart barcode" tags used to track packages and retail inventory could also detect changes in the real world—like temperature, pressure or weight—without batteries or added hardware?

That idea's now a reality, thanks to a team of researchers out of UC San Diego's Center for Wireless Communications (CWC) and Qualcomm Institute who have developed a robust and real-time RFID-based passive sensing system, i.e. a technique that can measure naturally occurring phenomena using harvested radio frequency (RF) energy.

Ishan Bansal, a second-year master's student in computer engineering, and Dinesh Bharadia, an associate professor of electrical and computer engineering at the Jacobs School of Engineering, are spearheading the charge with a new paper.

Named SenSync, the software-based innovation recently earned the Best Paper Award at the 2025 IEEE RFID Conference. The paper is titled "SenSync: Real-Time and Accurate Passive Sensing."

"Unlike a lot of conferences, IEEE RFID has a lot of industry involvement," explained Nagarjun Bhat, a co-author of the award-winning paper, "SenSync: Real-time and Accurate Passive Sensing," and a Ph.D. student in electrical engineering at the Jacobs School of Engineering.

"This is a space where there's a lot of commercializable potential, because companies take the ideas from here and use them to build a product that will reach thousands of people in the future."

Lead author Bansal explained that SenSync is a major leap forward for passive sensing, as it highlights how existing RFID infrastructure can be retooled to enable real-time, battery-free sensing. Low-cost RFID tags are already used in billions of products globally.

While previous work from Bhat and CWC laid out how RFID tags can be used to sense data from a hardware standpoint, SenSync is what brings the image into focus.

When using a differential sensing system that involves two RFID ICs connected to the same antenna, it's possible for them to interfere with each other. Instead of sending their data to the reader simultaneously, Bansal noted the data would arrive sequentially. That can cause slight differences in the data streams, which makes it difficult to determine the correct values.

"My algorithm puts the two divergent sequences together and makes sense out of it," said Bansal. "This enables us to sense clearly, repurposing RFID to get reliable sensor data without wires or batteries."

Solving long-standing challenges in passive RFID sensing

RFID systems have traditionally been optimized for identification, not sensing—they're what's used to track packages or checked baggage. While other similar systems like Bluetooth exist, RFID devices are much more mature and widespread. About 50 billion RFID tags are sold every year, Bansal said, compared to four to five billion Bluetooth devices.

But despite their availability, implementing RFID systems isn't always easy.

Commercial RFID readers are required to follow strict communications protocols that often include frequency hopping or sequential tag reading; in practice, this means that possibly distorted signals and timing mismatches have historically made sensing applications notoriously unreliable.

SenSync solves these problems using an algorithm called dynamic time warping (DTW), which was originally used in speech recognition.

So, how does it work? SenSync synchronizes—hence the name—data streams from multiple passive RFID tags and uses differential signal behaviors to get reliable readings.

"We don't want to change the basic data," Bansal explained, "We just want to match it well."

He noted that SenSync delivers five times greater sensory resolution and eight times the data throughput compared to previous passive RFID sensing methods, meaning it can process up to 500 data samples per second with sub-degree error rates, even in complex, dynamic environments.

"We were able to show that the data stream could be pushed every second without losing any accuracy or fidelity," he added.

Real-world sensing, reimagined

That improved sensing capability opens up practical applications ranging from automating warehouses and monitoring agricultural plots to improving medical sensing and measuring food waste.

Already, Bansal, Bhat and the rest of their research team have highlighted how SenSync can work in an augmented reality (AR) environment and were also recognized at the SenSys 2025 conference (co-hosted by ACM and IEEE) for having the Best Demo—Runner-up.

Using a smartphone-based AR app, conference attendees at IEEE RFID 2025 saw in real time how RFID tags could show the amount of force applied or visualize the moisture content in a pot of soil through AR.

Or, Bhat said, imagine a warehouse full of boxes upon boxes. You could just point your phone camera or AR glasses at a box and instantly see its weight without ever touching it, he explained.

"AR today is mostly limited to what your vision feed provides you," he noted. "But, the multimodal sensing enriches the data by pulling in the stimuli like temperature that we can't see. That's all done by the algorithm; without that, all the data you'd collect would be noisy and difficult to process in real time."

And, because SenSync is battery-free, runs on commercial RFID tags and requires no training data or recalibration, Bansal said that it offers a more sustainable, scalable sensing solution that can operate reliably in indoor, outdoor, or mobile environments.

"Our solution is actually truly ubiquitous," he said, explaining that "you can drop it anywhere, and it'll work the same way."

From RFID to physical AI

Bhat, Bansal and Bharadia feel that SenSync could be more than just an upgrade to passive sensing—it could be the foundation of physical AI that links the real and virtual worlds.

"Large language models (LLMs) are powering AI around us," Bharadia explained, adding that most machine learning draws on text, voice or images, all of which are widely available and easily accessible. But other sensory information like temperature and humidity aren't yet available for someone to build an LLM. "To empower our physical spaces with AI, you would need sensors and sensing."

"SenSync is truly an innovation," he added, "that can provide that battery-free sensing and do so with extremely low power usage and no wires."

More information: Ishan Bansal et al, SenSync: Real-Time and Accurate Passive Sensing, *2025 IEEE International Conference on RFID (RFID)* (2025). DOI: 10.1109/RFID64926.2025.11015540
Provided by University of California - San Diego

JUNE 12, 2025

Less is more: Efficient pruning for reducing AI memory and computational cost

by Bar-Ilan University edited by Gaby Clark, reviewed by Robert Egan

Single filter performance. Credit: *Physical Review E* (2025). DOI: 10.1103/49t8-mh9
Deep learning and AI systems have made great headway in recent years, especially in their capabilities of automating complex computational tasks such as image recognition, computer vision and natural language processing. Yet, these systems consist of billions of parameters and require great memory usage as well as expensive computational cost.

This reality raises the question: Can we optimize, or more correctly, prune, the parameters in those systems without compromising their capabilities? In a study just published in *Physical Review E* by researchers from Bar-Ilan University, the answer is a resounding yes.

In the article, the researchers show how a better understanding of the mechanism underlying successful deep learning leads to an efficient pruning of unnecessary parameters in a deep architecture without affecting its performance.

Researchers from Bar-Ilan University have developed a groundbreaking method to drastically reduce the size and energy consumption of deep learning systems—without compromising performance. Published in *Physical Review E*, their study reveals that by better understanding how deep networks learn, it's possible to prune up to 90% of parameters in certain layers while maintaining accuracy. This advancement, led by Prof. Ido Kanter and Ph.D. student Yarden Tzach, could make AI more efficient, sustainable, and scalable for real-world applications. Credit: Prof. Ido Kanter, Bar-Ilan University
"It all hinges on an initial understanding of what happens in deep networks, how they learn and what parameters are essential to its learning," said Prof. Ido Kanter, of Bar-Ilan's Department of Physics and Gonda (Goldschmied) Multidisciplinary Brain Research Center, who led the research.

"It's the ever-present reality of scientific research. The more we know, the better we understand, and in turn, the better and more efficient the technology we can create."

"There are many methods that attempt to improve memory and data usage," said Ph.D. student Yarden Tzach, a key contributor to this research.

"They were able to improve memory usage and computational complexity, but our method was able to prune up to 90% of the parameters of certain layers, without hindering the system's accuracy at all."

These results can lead to better usage of AI systems, both in memory as well as energy consumption. As AI becomes more and more prevalent in our day to day lives, reducing its energy cost will be of utmost importance.

More information: Yarden Tzach et al, Advanced deep architecture pruning using single-filter performance, *Physical Review E* (2025). DOI: 10.1103/49t8-mh9k. On *arXiv*: DOI: 10.48550/arxiv.2501.12880
Journal information: Physical Review E , arXiv
Provided by Bar-Ilan University

JUNE 12, 2025

Six ways AI can partner with us in creative inquiry, inspired by media theorist Marshall McLuhan

by Gordon A. Gow, The Conversation edited by Lisa Lock, reviewed by Andrew Zinin

Today's large language models (LLMs) process information across disciplines at unprecedented speed and are challenging higher education to rethink teaching, learning and disciplinary structures.

As AI tools disrupt conventional subject boundaries, educators face a dilemma: some seek to ban these tools, while others are seeking ways to embrace them in the classroom.

Both approaches risk missing a deeper transformation that was predicted 60 years ago by Canadian communication theorist Marshall McLuhan.

McLuhan's insights can help educators—and all of us grappling with the meaning, uses and misuses of AI—to think about how to cultivate a new mindset, one that integrates human agency and machine capabilities consciously and critically.

'Oracle of the electric age'

In the mid-1960s, McLuhan published *Understanding Media*, earning a reputation as the "oracle of the electric age."

In the chapter, "Automation: Learning a Living," McLuhan opens with a provocative observation: "Little Red Schoolhouse Dies When Good Road Built." Technological change, he suggested, doesn't merely augment existing systems—it transforms them.

While roads once expanded access to specialized education, automation reverses this logic, he argued.

This is because disciplinary boundaries are dissolved, and the intersection of learning and work is redefined. He wrote: "Automation … not only ends jobs in the world of work, it ends subjects in the world of learning."

McLuhan foresaw that computing would enable new forms of pattern recognition, requiring fundamentally different ways of thinking—more integrative, relational and responsive—rather than simply accelerating old methods.

Automation makes the arts mandatory

Crucially, McLuhan argued that far from making the liberal arts obsolete, automation makes them mandatory. In an age where machine intelligence is integrated into communication and creativity, the humanities, with their focus on cultural understanding, ethical reasoning and imaginative expression, become more essential than ever.

To navigate this landscape, we can borrow from complex systems researcher Stuart Kauffman's concept of the "adjacent possible," as developed in author and innovation expert Steven Johnson's theory of innovation.

The "adjacent possible" refers to the set of opportunities and innovations that become accessible when new combinations of existing ideas and technologies are explored.

This gives rise to what I refer to as AI-adjacency: a framework that treats artificial intelligence not as a replacement for human intelligence, but as a partner in strategic collaboration and creative inquiry.

Six ways AI can be a partner in creative inquiry

1. Critical discernment

AI-adjacent learning begins with critical discernment: the ability to assess intellectual and cultural value regardless of whether AI was involved in the creation process.

When game designer Jason Allen's AI-assisted image, *Théâtre D'opéra Spatial*, won first place in a digital arts competition at the 2022 Colorado State Fair—and Allen shared information about it on social media—controversy ensued.

Commenters were unsure how to evaluate artistic merit when creative direction is shared with AI. Allen reportedly spent more than 80 hours crafting more than 600 text prompts in Midjourney, and also digitally altered the work. The debate illustrates how critical discernment moves beyond detecting AI use to asking deeper questions about authorship, effort and aesthetic judgment.

2. Strategic collaboration

Strategic collaboration requires nuanced decision-making about when and how to involve AI tools in a creative process. A recent study reports that "the impact of ChatGPT as a feedback tool on students' writing skills was positive and significant."

As one student in the study noted: "When you use ChatGPT in a classroom, you're doing it with several people. So much talk going on simultaneously! It's kinda cool. The conversations are so meaningful and without noticing, we are working together and writing."

The value here is in an AI-facilitated collaboration that encourages students to become more interested in learning how to express themselves through writing.

3. Voice and vision stewardship

Stewarding voice and vision means ensuring that technology serves individual expression, not the other way around. At Berklee College of Music in Boston, with varied instructors, students are encouraged to explore AI's varied potential uses in enhancing their creative process. If it's used, instructors emphasize outputs must reflect the artist's own style, not just the algorithm's fluency. This fosters self-awareness and creative authorship amid technological collaboration.

4. Cultural and social responsibility

AI tools are not neutral, but they can be powerful allies when developed with cultural and social responsibility. Researchers on Vancouver Island are developing AI voice-to-text technology specifically for Kwak'wala, an endangered Indigenous language.

Sara Child, a Kwagu'ł band member and professor in Indigenous education leading the project, told CBC that by "building the technology tool, the speech recognition tool, we can tap into that amazing resource that will help us recapture and reclaim language that is trapped in archives."

Unlike existing systems designed for English, this AI must be built from scratch because Kwak'wala is verb-centered rather than noun-based.

The project demonstrates how AI can amplify marginalized voices. In this case, Indigenous communities control the development process and cultural knowledge remains in community hands.

5. Adaptive expertise

Adaptive expertise means knowing when to innovate beyond routine solutions. Medical education researchers Brian J. Hess and colleagues define it as "the capacity to apply not only routinized procedural approaches but also know when the situation calls for creative innovative solutions."

In an AI-integrated world, students must distinguish between when AI-generated responses are appropriate and can enhance productivity, versus when situations require human, slower, in-depth thinking and creative analysis.

For example, history students can use AI to quickly process archival materials and identify patterns, but must also learn how to use AI to help them interpret the cultural significance of those patterns, which requires innovative analytical approaches grounded in a liberal arts education.

6. Creative and intellectual agency

Creative and intellectual agency represents a central pillar of humanities education, rooted in the German concept of *Bildung*, which is developing oneself through critical engagement with complex ideas.

This principle of cultivating independent thinking and deep attention to challenging problems remains essential in an AI-integrated world. The challenge facing higher education is find ways to amplify intellectual agency through creative collaboration with AI tools. At Lehigh University in Pennsylvania, humanities students work with computer scientists to develop interdisciplinary courses like "Algorithms and Social Justice," which involves applying humanistic perspectives throughout data analysis processes.

McLuhan's warning: loss of self-awareness

McLuhan also offered a powerful warning through the myth of Narcissus in *Understanding Media*.

Contrary to popular view, McLuhan argued Narcissus didn't fall in love with himself; instead, he mistook his reflection for someone else.

This "extension of himself by mirror," McLuhan writes, "numbed his perceptions until he became the servomechanism of his own extended … image"—meaning, Narcissus became dependent on his own reflection.

The real danger of AI isn't replacement. It's the loss of self-awareness. We risk becoming passive users of our own technological extensions and allowing them to shape how we think, create and learn without realizing it. In McLuhan's terms, we become tools of our tools.

AI-adjacent practices offer a way out. By engaging consciously with technology through the six dimensions, students learn to use AI critically and creatively—without surrendering their agency.

Provided by The Conversation

JUNE 12, 2025

AI tools collect and store data about you from all your devices. Here's how to be aware of what you're revealing

by Christopher Ramezan, The Conversation edited by Lisa Lock, reviewed by Alexander Pol

Credit: CC0 Public Domain
Like it or not, artificial intelligence has become part of daily life. Many devices—including electric razors and toothbrushes—have become "AI-powered," using machine learning algorithms to track how a person uses the device, how the device is working in real time, and provide feedback. From asking questions to an AI assistant like ChatGPT or Microsoft Copilot to monitoring a daily fitness routine with a smartwatch, many people use an AI system or tool every day.

While AI tools and technologies can make life easier, they also raise important questions about data privacy. These systems often collect large amounts of data, sometimes without people even realizing their data is being collected. The information can then be used to identify personal habits and preferences, and even predict future behaviors by drawing inferences from the aggregated data.

As an assistant professor of cybersecurity at West Virginia University, I study how emerging technologies and various types of AI systems manage personal data and how we can build more secure, privacy-preserving systems for the future.

Generative AI software uses large amounts of training data to create new content such as text or images. Predictive AI uses data to forecast outcomes based on past behavior, such as how likely you are to hit your daily step goal, or what movies you may want to watch. Both types can be used to gather information about you.

How AI tools collect data

Generative AI assistants such as ChatGPT and Google Gemini collect all the information users type into a chat box. Every question, response and prompt that users enter is recorded, stored and analyzed to improve the AI model.

OpenAI's privacy policy informs users that "we may use content you provide us to improve our Services, for example to train the models that power ChatGPT." Even though OpenAI allows you to opt out of content use for model training, it still collects and retains your personal data. Although some companies promise that they anonymize this data, meaning they store it without naming the person who provided it, there is always a risk of data being reidentified.

Predictive AI

Beyond generative AI assistants, social media platforms like Facebook, Instagram and TikTok continuously gather data on their users to train predictive AI models. Every post, photo, video, like, share and comment, including the amount of time people spend looking at each of these, is collected as data points that are used to build digital data profiles for each person who uses the service.

The profiles can be used to refine the social media platform's AI recommender systems. They can also be sold to data brokers, who sell a person's data to other companies to, for instance, help develop targeted advertisements that align with that person's interests.

Many social media companies also track users across websites and applications by putting cookies and embedded tracking pixels on their computers. Cookies are small files that store information about who you are and what you clicked on while browsing a website.

One of the most common uses of cookies is in digital shopping carts: When you place an item in your cart, leave the website and return later, the item will still be in your cart because the cookie stored that information. Tracking pixels are invisible images or snippets of code embedded in websites that notify companies of your activity when you visit their page. This helps them track your behavior across the internet.

This is why users often see or hear advertisements that are related to their browsing and shopping habits on many of the unrelated websites they browse, and even when they are using different devices, including computers, phones and smart speakers. One study found that some websites can store over 300 tracking cookies on your computer or mobile phone.

Data privacy controls—and limitations

Like generative AI platforms, social media platforms offer privacy settings and opt-outs, but these give people limited control over how their personal data is aggregated and monetized. As media theorist Douglas Rushkoff argued in 2011, if the service is free, you are the product.

Many tools that include AI don't require a person to take any direct action for the tool to collect data about that person. Smart devices such as home speakers, fitness trackers and watches continually gather information through biometric sensors, voice recognition and location tracking. Smart home speakers continually listen for the command to activate or "wake up" the device. As the device is listening for this word, it picks up all the conversations happening around it, even though it does not seem to be active.

Some companies claim that voice data is only stored when the wake word—what you say to wake up the device—is detected. However, people have raised concerns about accidental recordings, especially because these devices are often connected to cloud services, which allow voice data to be stored, synced and shared across multiple devices such as your phone, smart speaker and tablet.

If the company allows, it's also possible for this data to be accessed by third parties, such as advertisers, data analytics firms or a law enforcement agency with a warrant.

Privacy rollbacks

This potential for third-party access also applies to smartwatches and fitness trackers, which monitor health metrics and user activity patterns. Companies that produce wearable fitness devices are not considered "covered entities" and so are not bound by the Health

Information Portability and Accountability Act. This means that they are legally allowed to sell health- and location-related data collected from their users.

Concerns about HIPAA data arose in 2018, when Strava, a fitness company, released a global heat map of users' exercise routes. In doing so, it accidentally revealed sensitive military locations across the globe through highlighting the exercise routes of military personnel.

The Trump administration has tapped Palantir, a company that specializes in using AI for data analytics, to collate and analyze data about Americans. Meanwhile, Palantir has announced a partnership with a company that runs self-checkout systems.

Such partnerships can expand corporate and government reach into everyday consumer behavior. This one could be used to create detailed personal profiles on Americans by linking their consumer habits with other personal data. This raises concerns about increased surveillance and loss of anonymity. It could allow citizens to be tracked and analyzed across multiple aspects of their lives without their knowledge or consent.

Some smart device companies are also rolling back privacy protections instead of strengthening them. Amazon recently announced that starting on March 28, 2025, all voice recordings from Amazon Echo devices would be sent to Amazon's cloud by default, and users will no longer have the option to turn this function off. This is different from previous settings, which allowed users to limit private data collection.

Changes like these raise concerns about how much control consumers have over their own data when using smart devices. Many privacy experts consider cloud storage of voice recordings a form of data collection, especially when used to improve algorithms or build user profiles, which has implications for data privacy laws designed to protect online privacy.

Implications for data privacy

All of this brings up serious privacy concerns for people and governments on how AI tools collect, store, use and transmit data. The biggest concern is transparency. People don't know what data is being collected, how the data is being used, and who has access to that data.

Companies tend to use complicated privacy policies filled with technical jargon to make it difficult for people to understand the terms of a service that they agree to. People also tend not to read terms of service documents. One study found that people averaged 73 seconds reading a terms of service document that had an average read time of 29-32 minutes.

Data collected by AI tools may initially reside with a company that you trust, but can easily be sold and given to a company that you don't trust.

AI tools, the companies in charge of them and the companies that have access to the data they collect can also be subject to cyberattacks and data breaches that can reveal sensitive personal information. These attacks can be carried out by cybercriminals who are in it for the money, or by so-called advanced persistent threats, which are typically nation/state-sponsored attackers who gain access to networks and systems and remain there undetected, collecting information and personal data to eventually cause disruption or harm.

While laws and regulations such as the General Data Protection Regulation in the European Union and the California Consumer Privacy Act aim to safeguard user data, AI development and use have often outpaced the legislative process. The laws are still catching up on AI and data privacy. For now, you should assume any AI-powered device or platform is collecting data on your inputs, behaviors and patterns.

Using AI tools

Although AI tools collect people's data, and the way this accumulation of data affects people's data privacy is concerning, the tools can also be useful. AI-powered applications can streamline workflows, automate repetitive tasks and provide valuable insights.

But it's crucial to approach these tools with awareness and caution.

When using a generative AI platform that gives you answers to questions you type in a prompt, don't include any personally identifiable information, including names, birth dates, Social Security numbers or home addresses. At the workplace, don't include trade secrets or classified information. In general, don't put anything into a prompt that you wouldn't feel comfortable revealing to the public or seeing on a billboard. Remember, once you hit enter on the prompt, you've lost control of that information.

Remember that devices which are turned on are always listening—even if they're asleep. If you use smart home or embedded devices, turn them off when you need to have a private conversation. A device that's asleep looks inactive, but it is still powered on and listening for a wake word or signal. Unplugging a device or removing its batteries is a good way of making sure the device is truly off.

Finally, be aware of the terms of service and data collection policies of the devices and platforms that you are using. You might be surprised by what you've already agreed to.

Provided by The Conversation

JUNE 12, 2025

AI literacy: What it is, what it isn't, who needs it, and why it's hard to define

by Daniel S. Schiff, Arne Bewersdorff, Marie Hornberger, The Conversation edited by Lisa Lock, reviewed by Alexander Pol

It is "the policy of the United States to promote AI literacy and proficiency among Americans," reads an executive order President Donald Trump issued on April 23, 2025. The executive order, titled Advancing Artificial Intelligence Education for American Youth, signals that advancing AI literacy is now an official national priority.

This raises a series of important questions: What exactly is AI literacy, who needs it, and how do you go about building it thoughtfully and responsibly?

The implications of AI literacy, or lack thereof, are far-reaching. They extend beyond national ambitions to remain "a global leader in this technological revolution" or even prepare an "AI-skilled workforce," as the executive order states. Without basic literacy, citizens and consumers are not well equipped to understand the algorithmic platforms and decisions that affect so many domains of their lives: government services, privacy, lending, health care, news recommendations and more. And the lack of AI literacy risks ceding important aspects of society's future to a handful of multinational companies.

How, then, can institutions help people understand and use—or resist—AI as individuals, workers, parents, innovators, job seekers, students, employers and citizens? We are a policy scientist and two educational researchers who study AI literacy, and we explore these issues in our research.

What AI literacy is and isn't

At its foundation, AI literacy includes a mix of knowledge, skills and attitudes that are technical, social and ethical in nature. According to one prominent definition, AI literacy refers to "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace."

AI literacy is not simply programming or the mechanics of neural networks, and it is certainly not just prompt engineering—that is, the act of carefully writing prompts for chatbots. Vibe coding, or using AI to write software code, might be fun and important, but restricting the definition of literacy to the newest trend or the latest need of employers won't cover the bases in the long term. And while a single master definition may not be needed, or even desirable, too much variation makes it tricky to decide on organizational, educational or policy strategies.

Who needs AI literacy? Everyone, including the employees and students using it, and the citizens grappling with its growing impacts. Every sector and sphere of society is now involved with AI, even if this isn't always easy for people to see.

Exactly how much literacy everyone needs and how to get there is a much tougher question. Are a few quick HR training sessions enough, or do we need to embed AI across K-12 curricula and deliver university micro credentials and hands-on workshops? There is much that researchers don't know, which leads to the need to measure AI literacy and the effectiveness of different training approaches.

Measuring AI literacy

While there is a growing and bipartisan consensus that AI literacy matters, there's much less consensus on how to actually understand people's AI literacy levels. Researchers have focused on different aspects, such as technical or ethical skills, or on different populations—for example, business managers and students—or even on subdomains like generative AI.

A recent review study identified more than a dozen questionnaires designed to measure AI literacy, the vast majority of which rely on self-reported responses to questions and statements such as "I feel confident about using AI." There's also a lack of testing to see whether these questionnaires work well for people from different cultural backgrounds.

Moreover, the rise of generative AI has exposed gaps and challenges: Is it possible to create a stable way to measure AI literacy when AI is itself so dynamic?

In our research collaboration, we've tried to help address some of these problems. In particular, we've focused on creating objective knowledge assessments, such as multiple-choice surveys tested with thorough statistical analyses to ensure that they accurately measure AI literacy. We've so far tested a multiple-choice survey in the U.S., U.K. and Germany and found that it works consistently and fairly across these three countries.

There's a lot more work to do to create reliable and feasible testing approaches. But going forward, just asking people to self-report their AI literacy probably isn't enough to understand where different groups of people are and what supports they need.

Approaches to building AI literacy

Governments, universities and industry are trying to advance AI literacy.

Finland launched the Elements of AI series in 2018 with the hope of educating its general public on AI. Estonia's AI Leap initiative partners with Anthropic and OpenAI to provide access to AI tools for tens of thousands of students and thousands of teachers. And China is now requiring at least eight hours of AI education annually as early as elementary school, which goes a step beyond the new U.S. executive order. On the university level, Purdue University and the University of Pennsylvania have launched new master's in AI programs, targeting future AI leaders.

Despite these efforts, these initiatives face an unclear and evolving understanding of AI literacy. They also face challenges to measuring effectiveness and minimal knowledge on what teaching approaches actually work. And there are long-standing issues with respect to equity—for example, reaching schools, communities, segments of the population and businesses that are stretched or under-resourced.

Next moves on AI literacy

Based on our research, experience as educators and collaboration with policymakers and technology companies, we think a few steps might be prudent.

Building AI literacy starts with recognizing it's not just about tech: People also need to grasp the social and ethical sides of the technology. To see whether we're getting there, we researchers and educators should use clear, reliable tests that track progress for different age groups and communities. Universities and companies can try out new teaching ideas first, then share what works through an independent hub. Educators, meanwhile, need proper training and resources, not just additional curricula, to bring AI into the classroom. And because opportunity isn't spread evenly, partnerships that reach under-resourced schools and neighborhoods are essential so everyone can benefit.

Critically, achieving widespread AI literacy may be even harder than building digital and media literacy, so getting there will require serious investment—not cuts—to education and research. There is widespread consensus that AI literacy is important, whether to boost AI trust and adoption or to empower citizens to challenge AI or shape its future. As with AI itself, we believe it's important to approach AI literacy carefully, avoiding hype or an overly technical focus. The right approach can prepare students to become "active and responsible participants in the workforce of the future" and empower Americans to "thrive in an increasingly digital society," as the AI literacy executive order calls for.

Provided by The Conversation

JUNE 12, 2025

Q&A: Why improving robot design is essential to achieving true intelligence

by Rimma Gerenstein, Albert Ludwig University of Freiburg

edited by Lisa Lock, reviewed by Robert Egan

Jun.-Prof. Edoardo Milana, Junior Professor for Soft Machines at the University of Freiburg. Credit: Lucia Brunold

Thanks to artificial intelligence, robots can already perform many tasks that would otherwise require humans. In this interview, Edoardo Milana, a junior professor of soft machines in the Department of Microsystems Engineering at the University of Freiburg, explains how improved design and innovative mechanics are broadening the range of applications for these machines.

Why is there a need for an alternative to conventional robots?

Robots can already perform amazing things with the help of artificial intelligence and machine learning. However, all this intelligence is concentrated in the software—the brain—and no comparable focus is placed on the mechanical design—the body. As such, robots are pretty much like puppets. Software is used to try to exert full control over all the body's movements.

This approach requires the hardware to be very simple from a mechanical point of view and easier to operate using digital microcontrollers. Depending on the area of application, this may be sufficient and even needed to meet the precision and high force requirements. However, when we look at motion efficiency and agility, the performance of robots falls far short of that of living beings.

Nevertheless, there are already robots that imitate animals, such as dogs and cats

These quadrupeds—and even humanoids—are impressive engineering masterpieces, however, they cannot compete with real animals in terms of motion agility. They also consume a lot of energy to move, whereas animals and humans can perform much more complex movements using far less energy. A quadruped walking at normal pace consumes roughly 300 Watt on average to drive its 12 motors, the robot's "muscles," while a dog consumes 30 Watt to activate hundreds of muscles.

This is possible because, in nature, movements strongly rely on the mechanical properties of the body, which passively and actively adapts to the external forces exerted by the environment, harnessing the compliance of biological materials. Beyond digital control, the focus in robotics should also be on implementing intelligence, or "embodied intelligence," into the design of the robot. This would free up computing capacity and energy currently used for low-level motion control for the high-level logical operations of the robot, such as reasoning, planning and perceiving.

The concept of embodied intelligence originates from the fields of philosophy and psychology. But what does it mean for you as an engineer who develops robots?

For me, the interesting thing is that the theory behind it can be applied not only to biological beings, but to robots, too. The basic idea is that physical interaction between the body and the environment shapes intelligent behavior. It's not just about having a body controlled by the mind—this control lies partly in the body itself and in the way it interacts with the mind.

In robotics, this means that if we want a truly intelligent robot, we can't just build a body consisting of two or three metal bars and a few joints, then put a very intelligent computer inside it. If that were the case, we would already have robots with completely different capabilities.

What could such intelligent robots look like instead?

I am researching soft robots made of soft materials, which could be considered as inspired by primitive and aquatic biological organisms. There are already robots in this field whose control is based entirely on physical principles, and that do not require digital microcontrollers. They utilize the non-linear physical properties of soft materials to generate the control signals that drive the robot.

Together with researchers from Stuttgart, the Netherlands, and Belgium, I have written an article published in *Science Robotics* presenting such soft robots, which introduces a new concept: The concept of "physical control." We have identified three particular control mechanisms for such soft robots.

One interesting example is robots with self-oscillating valves. When air pressure is added, the valves open and close again, increasing and then releasing air pressure. This transmits a rhythmic air pressure signal through the system, controlling the movement of the individual robot parts.

In the future, we will need to find a compromise: We won't be able to manage without software and microcontrollers in robotics, but we can achieve a lot through better robot body design.

More information: Edoardo Milana et al, Physical control: A new avenue to achieve intelligence in soft robotics, *Science Robotics* (2025). DOI: 10.1126/scirobotics.adw7660
Journal information: Science Robotics
Provided by Albert Ludwig University of Freiburg

JUNE 12, 2025

Cellular coordinate system reveals secrets of active matter

by Lori Dajose, California Institute of Technology

edited by Lisa Lock, reviewed by Robert Egan

Left: Active matter composed of filaments and motors. Center: Active matter overlaid with a fluorescence-cancelling grid, creating a coordinate system to measure deformation. Right: As the system contracts, the coordinate system deforms as well. Credit: S. Hirokawa
All humans who have ever lived were once each an individual cell, which then divided countless times to produce a body made up of about 10 trillion cells. These cells have busy lives, executing all kinds of dynamic movement: contracting every time we flex a muscle, migrating toward the site of an injury, and rhythmically beating for decades on end.

Cells are an example of active matter. As inanimate matter must burn fuel to move, like airplanes and cars, active matter is similarly animated by its consumption of energy. The basic molecule of cellular energy is adenosine triphosphate (ATP), which catalyzes chemical reactions that enable cellular machinery to work.

Caltech researchers have now developed a bioengineered coordinate system to observe the movement of cellular machinery. The research enables a better understanding of how cells create order out of chaos, such as during embryonic development or in the organized movements of chromosomes that are a prerequisite to faithful cell division.

The work was conducted in the laboratories of Rob Phillips, the Fred and Nancy Morris Professor of Biophysics, Biology, and Physics, and Matt Thomson, Professor of Computational Biology and Heritage Medical Research Institute Investigator. A paper describing the study is published in the *Proceedings of the National Academy of Sciences.*

The basic units of cellular machinery are motors and filaments made of proteins, which act like the muscles and skeleton of the cell. These structures self-assemble, like little protein robots, to enable cells to move. In 2018, former graduate student Tyler Ross (Ph.D. '21) engineered a system of these components that can be controlled by light in a lab setting, enabling researchers to observe and experiment upon their movements. Each experimental system is only the width of a human hair, containing thousands of individual motors and filaments.

An example of bioengineered microtubules being directed by light. Credit: Caltech
In the new work, led by former graduate student Soichi Hirokawa (Ph.D. '23), the team developed additional light patterns that create a grid, or coordinate system, upon the mixture of motors and filaments. To understand this, imagine a sheet of rubber with a grid patterned on it—as the rubber stretches and deforms, the grid does as well.

Once a set of regularly spaced squares, the grid's deformation gives a measure of which regions are being stretched or squeezed and by how much. In this way, the team can track the movements of a collection of filaments and motors—they are too small to be seen themselves, but the light-patterned grid, each square about 12-by-12 micrometers, is visible with a microscope.

"The system allows us to observe how these biomolecules reorganize as they collectively form a structure," says Hirokawa. "With it, we can distinguish the processes that contribute to the deformations that we observe on these squares."

This new system enabled the team to measure the competing dynamics of active shrinking and a process that influences cellular self-assembly, called diffusion. Taking a mixture of motors and filaments, the researchers triggered the components to contract inward, like a shrinking circle. But each component naturally still experiences some random movement, or diffusion, jiggling every which way as the whole contracts.

The deforming coordinate system enabled the team to watch this competition between active contraction and random diffusion, and characterize it. Interestingly, they found that the more ATP is in the system, the more the molecules randomly diffuse.

In 2019, Caltech researchers demonstrated how to control cellular active matter with light. Credit: California Institute of Technology
"The formation of patterns and structure in biology has to fight against this randomness," says Phillips. "The system is able to organize despite the forces of chaos."

The dynamic coordinate system introduced here could be used in other contexts as well.

"Order is particularly important in processes like embryonic development," says staff scientist and co-author Heun Jin Lee. "An early embryo gastrulates, folding into a tube that ultimately becomes the digestive tract. You could imagine decorating the surface of an embryo with a coordinate system that stretches as the embryo folds."

More information: Soichi Hirokawa et al, Motor-driven microtubule diffusion in a photobleached dynamical coordinate system, *Proceedings of the National Academy of Sciences* (2025). DOI: 10.1073/pnas.2417020122
Journal information: Proceedings of the National Academy of Sciences

---

JUNE 12, 2025

Unprecedented dataset of molecular simulations to train AI models released

by Los Alamos National Laboratory

edited by Sadie Harley, reviewed by Robert Egan

In this representation made with the Architector software and an example from the Open Molecules 2025 dataset, lanthanum, a rare earth metal, is surrounded by diverse bonding molecules. Lanthanum alloys are used in batteries and hydrogen gas applications. Credit: Los Alamos National Laboratory
A collaborative effort between Meta, Lawrence Berkeley National Laboratory and Los Alamos National Laboratory leverages Los Alamos' expertise in building tools for molecular

screening capabilities. The release of "Open Molecules 2025", an unprecedented dataset of molecular simulations, can accelerate opportunities for machine learning to transform research in fields such as biology, materials science and energy technologies.

The dataset appears on the *arXiv* preprint server.

"A prohibitive part of molecular design has been the extreme computational cost needed to achieve quantum chemistry-level accuracy," said Michael G. Taylor, researcher at Los Alamos and project member.

"In order to train machine learning models capable of quantum chemistry-level accuracy, we need vast amounts of diverse, valid training data. Open Molecules 2025 bridges this gap with a dataset of over 100 million density-functional theory calculations that we can use to train machine learning models accurately enough for all kinds of chemical challenges."

The dataset is key to unlocking the use of machine learning potentials for chemical applications, such as designing a new drug to fight disease or a battery cell to store energy.

The employment of density functional theory calculations in the dataset enables a precise, atomic-level understanding of molecular behavior and interactions. Unique software designed by Taylor played a critical role in the ability of Open Molecules 2025 to reach its goals.

Novel software helps build the dataset

To help run the calculations and build the dataset, the collaboration leveraged the capabilities of the Architector software, designed by Taylor. Architector is a state-of-the-art software for predicting 3D structures of metal complexes.

Metal complexes are chemicals in which a central metal atom is bound to an array of other molecules or atoms, and they represent important chemistry relevant to applications from biology to materials science.

Architector, as employed by Taylor and collaborators in the Lab's Theoretical division, has mainly been applied to "F-block" elements: lanthanides like cerium and ytterbium, and actinides such as thorium and uranium.

The F-block elements include many elements often referred to as rare earth elements, which are valuable for an array of industrial purposes, including high-tech applications in telecommunications, imaging, data storage and more.

The metal complexes represent an important class of chemistry explored with the Open Molecules 2025 dataset. Other classes include ion molecules such as proteins and RNA, small molecules that might be the basis of drug discovery, and electrolyte metals surrounded by different solvents. Taylor estimates that the chemistry explored by Architector represents up to a third of the entire dataset.

An investment in foundational chemistry knowledge

Meta tasked its vast computing power to run the density functional theory calculations. Considering only the rare earth molecular simulations it was able to achieve, the Open Molecules 2025 project resulted in data on approximately 20,000 structures on each of the 17 rare earth elements.

The next-largest dataset available in literature has approximately 1,000 structures total per rare earth element.

The immense data generated can now be used to train other machine learning models at a fraction of the time and cost. The dataset could lead to pre-trained foundation models that can be fine-tuned with minimal added data in areas of interest.

The entire Open Molecules 2025 effort, including initial machine learning models trained on the data, will be open to the public, giving researchers the ability to use data and models relevant to their research.

"Chemical design often boils down to predicting the properties of new chemistries with minimal information and computational expense," said Taylor.

"Having this dataset, with the ability to train machine learning models to do that predictive work, is potentially transformative for scientific discovery."

More information: Daniel S. Levine et al, The Open Molecules 2025 (OMol25) Dataset, Evaluations, and Models, *arXiv* (2025). DOI: 10.48550/arxiv.2505.08762
Journal information: arXiv
Provided by Los Alamos National Laboratory

JUNE 13, 2025

Rethinking AI: Researchers propose a more effective, human-like approach

by Rensselaer Polytechnic Institute edited by Sadie Harley, reviewed by Robert Egan

Future of artificial neural networks. Credit: *Patterns* (2025). DOI: 10.1016/j.patter.2025.101231
New research from Rensselaer Polytechnic Institute (RPI) could help shape the future of artificial intelligence by making AI systems less resource-intensive, higher performing, and

designed to emulate the human brain. The research was published in *Patterns*, titled "Dimensionality and dynamics for next-generation neural networks."

As AI models grow ever larger, so do their costs and limitations. Researchers at RPI and City University of Hong Kong offer a potential solution: instead of expanding outward with more layers and data, they propose building upward—adding internal structure that mirrors a 3D biological neural network and incorporating recursive loops to enhance network introspection.

This vertical dimension and loop allow artificial neural networks to process information more effectively and efficiently, potentially transforming AI's ability to learn and adapt in 3D and higher dimensions.

"This new AI framework not only boosts efficiency but also unlocks practical opportunities," said Wang. "This research could be a crucial step toward driving advancements in next-generation artificial neural networks, closely relevant to health care and education, while paving the way for deeper insights into how the human brain works."

By introducing a vertical "height" dimension and feedback loops that allow artificial neural networks to relate, reflect, and refine outputs, the new design ideas could make AI models much smarter, leading to potentially major implications.

Doing more with fewer resources could help expand access to advanced AI technologies, reduce the environmental footprint of massive model training, and enable more real-time applications—from robotics to personalized medicine.

One such application lies in neuroscience, where brain-inspired neural networks could help scientists better understand cognition and even uncover new clues about neurological disorders such as Alzheimer's and epilepsy.

"This framework isn't just about smarter AI—it's about more sustainable, accessible, and explainable AI," Wang said. "And it may help us learn more about our own brains along the way."

The collaborative study was led by Ge Wang, Ph.D., Clark & Crossan Endowed Chair and director of the Biomedical Imaging Center at Rensselaer, and Fenglei Fan, Ph.D., Wang's former Ph.D. student and current assistant professor at City University of Hong Kong.

This work builds on RPI's longstanding leadership in AI research. Through major initiatives such as the AI Research Collaboration with IBM and the Future of Computing Institute, RPI researchers are developing cutting-edge AI technologies that are designed not only to meet human needs, but to redefine them.

From advancing human-machine collaboration to exploring brain-inspired computing, RPI is committed to shaping the future of AI.

More information: Ge Wang et al, Dimensionality and dynamics for next-generation artificial neural networks, *Patterns* (2025). DOI: 10.1016/j.patter.2025.101231

Journal information: Patterns
Provided by Rensselaer Polytechnic Institute

Top AI Researchers Meet to Discuss What Comes After Humanity

"AGI is likely to end humanity." JUN 16

A group of the top minds in AI gathered over the weekend to discuss the "posthuman transition" — a mind-bending exercise in imagining a future in which humanity willfully hands over power, or perhaps bequeaths existence entirely, to some sort of superhuman intelligence.

As *Wired* reports, the lavish party was organized by generative AI entrepreneur Daniel Faggella. Attendees included "AI founders from $100 million to $5 billion valuations" and "most of the important philosophical thinkers on AGI," Faggella enthused in a LinkedIn post.

He organized the soirée at a $30 million mansion in San Francisco because the "big labs, the people that know that AGI is likely to end humanity, don't talk about it because the incentives don't permit it," Faggella told *Wired*.

The symposium allowed attendees and speakers alike to steep themselves in a largely fantastical vision of a future where artificial general intelligence (AGI) was a given, rather than some distant dream of tech that isn't even close to existing.

AI companies, most notably OpenAI, have talked at length about wanting to realize AGI, though often without clearly defining the term.

The risks of racing toward a superhuman intelligence have remained hotly debated, with billionaire Elon Musk once arguing that unregulated AI could be the "biggest risk we face as a civilization." OpenAI Sam Altman has also warned of dangers facing humanity, including increased inequality and population control through mass surveillance, as a result of realizing AGI — which also happens to be his firm's number one priority.

But for now, those are largely moot points made by individuals who are billions of dollars deep in reassuring investors that AGI is mere years away. Given the current state of wildly hallucinating large language models that still fail at the most basic tasks, we are seemingly still a long way from a point at which AI could surpass the intellectual capabilities of humans.

Just last week, researchers at Apple released a damning paper that threw cold water on the "reasoning" capabilities of the latest and most powerful LLMs, arguing they "face a complete accuracy collapse beyond certain complexities."

However, to insiders and believers in the tech, AGI is mostly a matter of when, not if. Speakers at this weekend's event talked about how AI can seek out deeper, universal values that humanity hasn't even been privy to, and that machines should be taught to pursue "the good," or risk enslaving an entity capable of suffering.

As *Wired* reports, Faggella similarly invoked philosophers including Baruch Spinoza and Friedrich Nietzsche, calling on humanity to seek out the yet-undiscovered value in the universe.

"This is an advocacy group for the slowing down of AI progress, if anything, to make sure we're going in the right direction," he told the publication.

Honest AI

AI lets the dead speak again—in court, in memory, and beyond.

AI identifies key gene sets that cause complex diseases

Northwestern University researchers develop AI tool to uncover gene networks, paving the way for personalised treatments.

From Northwestern University 14/06/25 (first released 09/06/25)

Northwestern University biophysicists have developed a new computational tool for identifying the gene combinations underlying complex illnesses like diabetes, cancer and asthma. Unlike single-gene disorders, these conditions are influenced by a network of multiple genes working together. Credit: Camila Felix

Northwestern University biophysicists have developed a new computational tool for identifying the gene combinations underlying complex illnesses like diabetes, cancer and asthma.

Unlike single-gene disorders, these conditions are influenced by a network of multiple genes working together.

But the sheer number of possible gene combinations is huge, making it incredibly difficult for researchers to pinpoint the specific ones that cause disease.

Using a generative artificial intelligence (AI) model, the new method amplifies limited gene expression data, enabling researchers to resolve patterns of gene activity that cause complex traits.

This information could lead to new and more effective disease treatments involving molecular targets associated with multiple genes.

The study will be published during the week of June 9 in the Proceedings of the National Academy of Sciences.

"Many diseases are determined by a combination of genes — not just one," said Northwestern's Adilson Motter, the study's senior author.

"You can compare a disease like cancer to an airplane crash.

In most cases, multiple failures need to occur for a plane to crash, and different combinations of failures can lead to similar outcomes.

This complicates the task of pinpointing the causes.

Our model helps simplify things by identifying the key players and their collective influence."

An expert on complex systems, Motter is the Charles E. and Emma H. Morrison Professor of Physics at Northwestern's Weinberg College of Arts and Sciences and the director of the Center for Network Dynamics.

The other authors of the study — all associated with Motter's Lab — are postdoctoral researcher Benjamin Kuznets-Speck, graduate student Buduka Ogonor and research associate Thomas Wytock.

Current methods fall short

For decades, researchers have struggled to unravel the genetic underpinnings of complex human traits and diseases.

Even non-disease traits like height, intelligence and hair color depend on collections of genes.

Existing methods, such as genome-wide association studies, try to find individual genes linked to a trait.

But they lack the statistical power to detect the collective effects of groups of genes.

“The Human Genome Project showed us that we only have six times as many genes as a single-cell bacterium,” Motter said.

“But humans are much more sophisticated than bacteria, and the number of genes alone does not explain that.

This highlights the prevalence of multigenic relationships, and that it must be the interactions among genes that give rise to complex life.”

“Identifying single genes is still valuable,” Wytock added.

“But there is only a very small fraction of observable traits, or phenotypes, that can be explained by changes in single genes.

Instead, we know that phenotypes are the result of many genes working together.

Thus, it makes sense that multiple genes typically contribute to the variation of a trait.”

Not genes but gene expression

To help bridge the long-standing knowledge gap between genetic makeup (genotype) and observable traits (phenotype), the research team developed a sophisticated approach that combines machine learning with optimization.

Called the Transcriptome-Wide conditional Variational auto-Encoder (TWAVE), the model leverages generative AI to identify patterns from limited gene expression data in humans.

Accordingly, it can emulate diseased and healthy states so that changes in gene expression can be matched with changes in phenotype.

Instead of examining the effects of individual genes in isolation, the model identifies groups of genes that collectively cause a complex trait to emerge.

The method then uses an optimization framework to pinpoint specific gene changes that are most likely to shift a cell’s state from healthy to diseased or vice versa.

“We’re not looking at gene sequence but gene expression,” Wytock said.

“We trained our model on data from clinical trials, so we know which expression profiles are healthy or diseased.

For a smaller number of genes, we also have experimental data that tells how the network responds when the gene is turned on or off, which we can match with the expression data to find the genes implicated in the disease."

Focusing on gene expression has multiple benefits.

First, it bypasses patient privacy issues.

Genetic data — a person's actual DNA sequence — is inherently unique to an individual, providing a highly personal blueprint of health, genetic predispositions and family relationships.

Expression data, on the other hand, is more like a dynamic snapshot of cellular activity.

Second, gene expression data implicitly accounts for environmental factors, which can turn genes "up" or "down" to perform various functions.

"Environmental factors might not affect DNA, but they definitely affect gene expression," Motter said.

"So, our model has the benefit of indirectly accounting for environmental factors."

A path to personalized treatment

To demonstrate TWAVE's effectiveness, the team tested it across several complex diseases.

The method successfully identified the genes — some of which were missed by existing methods — that caused those diseases.

TWAVE also revealed that different sets of genes can cause the same complex disease in different people.

That finding suggests personalized treatments could be tailored to a patient's specific genetic drivers of disease.

"A disease can manifest similarly in two different individuals," Motter said.

"But, in principle, there could be a different set of genes involved for each person owing to genetic, environmental and lifestyle differences.

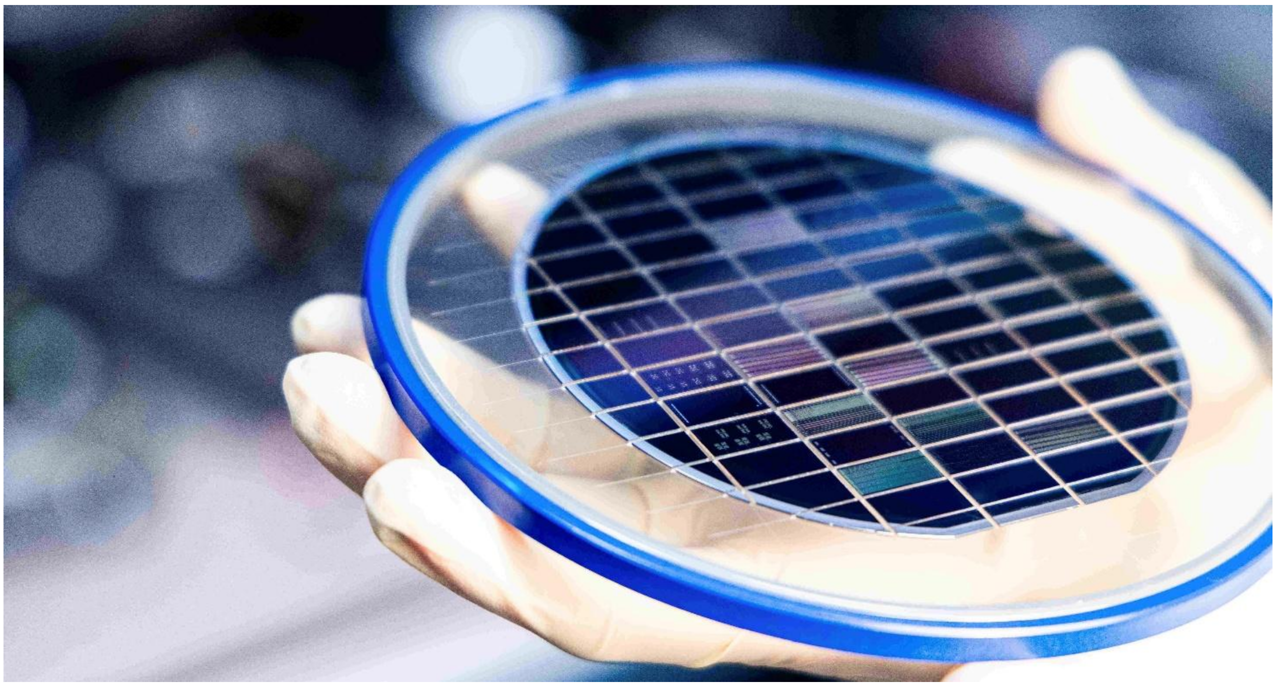
This information could orient personalized treatment."

The study, "Generative prediction of causal gene sets responsible for complex traits," was supported by the National Cancer Institute (grant number P50-CA221747) through the Malnati Brain Tumor Institute, the NSF-Simons National Institute for Theory and Mathematics in Biology (National Science Foundation grant number DMS-2235451 and Simons Foundation *grant number MP-TMPS-00005320) and the National Science Foundation (grant number MCB-2206974).*

World's first light-powered neural processing units (NPUs) could massively reduce energy consumption in AI data centers

News By Owen Hughes published April 3, 2025

Q.ANT's new chip uses photon power in a bid to solve AI's big energy issue. It's also 50 times faster than silicon-based equivalents, the company says.



The Q.ANT wafer based on Thin Film Lithium Niobate enables photonic integrated circuits with high-precision, high-speed optical modulation, low noise and reduced thermal dissipation. (Image credit: © Q.ANT)

A light-powered computer chip designed to drive artificial intelligence (AI) data centers and make high-performance computing (HPC) more sustainable has entered production. In a statement published Feb. 24, representatives from analog photonic chip company Q.ANT said its photonic AI chip could deliver a 30-fold increase in energy efficiency and a 50-fold boost in computing speed compared with conventional, silicon-based computer chips.

Pilot production of the new chip is now underway at IMS Chips in Stuttgart, Germany, where Q.ANT has invested 14 million euros ($15.1 million) to repurpose an existing semiconductor factory to fabricate its new, light-powered chip.

Because the chip is being produced on a repurposed facility instead of a specialist production line, the company believes it can bring the technology to market much more quickly. The chip can also integrate with the existing HPC servers, potentially accelerating adoption, Q.ANT representatives said.

"By 2030, we aim to make our photonic processors a scalable, energy-efficient cornerstone of AI infrastructure," Michael Förtsch, chief executive of Q.ANT, said in the statement.

Photonic computing

Photonic chips could solve a massive challenge faced by existing processor technology, particularly as AI and other data- and resource-intensive computing applications grow.

Traditional silicon chips control electrical signals using tiny switches called transistors. Photonic chips, by contrast, process data using light particles (photons), which are massless and can travel much faster than electrons do in conventional computer chips. Photons don't emit heat in the same way electrons carrying an electrical charge do. As such, using photonic chips in applications involving complex, energy-intensive computations like AI could overcome the limitations of classic silicon chip architecture and thus vastly accelerate the computers' processing speed and reduce their energy consumption.

"This comes at a critical time for the computing industry, as the exponential growth of AI and data-intensive applications will soon overwhelm the current data center infrastructure," Jens Anders, a professor at the University of Stuttgart and director and chief executive of IMS Chips, said in the statement. Anders added that the two companies aimed to establish "a scalable model for energy-efficient computing."

Q.ANT's chip is built using thin-film lithium niobate (TFLN), a crystalline compound applied to a wafer that forms the basis of the company's photonic chip. TFLN is increasingly catching the attention of photonics researchers and quantum scientists for its potential in next-generation computing. When an electric field is applied to the material, it can be used to control the speed and phase of light waves, thereby enabling it to modulate optical signals with extreme precision.

The pilot production line has been set up specifically to manufacture chips that incorporate TFLN, with Q.ANT aiming to fabricate 1,000 wafers per year.

"As AI and data-intensive applications push conventional semiconductor technology to its limits, we need to rethink the way we approach computing at the core," Förtsch said. "With this pilot line, we are accelerating time to market and laying the foundation for photonic processors to become standard coprocessors in high-performance computing."

Brain study reveals how humans intuitively navigate different environments, offering direction for better AI

by University of Amsterdam edited by Gaby Clark, reviewed by Robert Egan

Two-dimensional multidimensional scaling of locomotive action affordance space. Credit: *Proceedings of the National Academy of Sciences* (2025). DOI: 10.1073/pnas.2414005122

How do you intuitively know that you can walk on a footpath and swim in a lake? Researchers from the University of Amsterdam have discovered unique brain activations that reflect how we can move our bodies through an environment.

Published in *Proceedings of the National Academy of Sciences*, the study not only sheds new light on how the human brain works, but also shows where artificial intelligence is lagging behind. According to the researchers, AI could become more sustainable and human-friendly if it incorporated this knowledge about the human brain.

When we see a picture of an unfamiliar environment—a mountain path, a busy street, or a river—we immediately know how we could move around in it: walk, cycle, swim or not go any further. That sounds simple, but how does your brain actually determine these action opportunities?

Ph.D. student Clemens Bartnik and a team of co-authors show how we make estimates of possible actions thanks to unique brain patterns. The team, led by computational neuroscientist Iris Groen, also compared this human ability with a large number of AI models, including ChatGPT.

"AI models turned out to be less good at this and still have a lot to learn from the efficient human brain," Groen concludes.

Viewing images in the MRI scanner

Using an MRI scanner, the team investigated what happens in the brain when people look at various photos of indoor and outdoor environments. The participants used a button to indicate whether the image invited them to walk, cycle, drive, swim, boat or climb. At the same time, their brain activity was measured.

"We wanted to know: when you look at a scene, do you mainly see what is there—such as objects or colors—or do you also automatically see what you can do with it," says Groen.

"Psychologists call the latter 'affordances'—opportunities for action; imagine a staircase that you can climb, or an open field that you can run through."

Scene-selective ROIs. Scene-selective ROI (Parahippocampal Place Area, PPA; Occipital Place Area, OPA; and Medial Place Area, MPA) masks of subject 004 in MNI-152 template space (see SI Methods for details). Credit: *Proceedings of the National Academy of Sciences* (2025). DOI: 10.1073/pnas.2414005122

Unique processes in the brain

The team discovered that certain areas in the visual cortex become active in a way that cannot be explained by visible objects in the image. "What we saw was unique," says Groen. "These brain areas not only represent what can be seen, but also what you can do with it."

The brain did this even when participants were not given an explicit action instruction. "These action possibilities are therefore processed automatically," says Groen. "Even if you do not consciously think about what you can do in an environment, your brain still registers it."

The research thus demonstrates for the first time that affordances are not only a psychological concept, but also a measurable property of our brains.

What AI doesn't understand yet

The team also compared how well AI algorithms—such as image recognition models or GPT-4—can estimate what you can do in a given environment. They were worse at predicting possible actions. "When trained specifically for action recognition, they could somewhat approximate human judgments, but the human brain patterns didn't match the models' internal calculations," Groen explains.

"Even the best AI models don't give exactly the same answers as humans, even though it's such a simple task for us," Groen says. "This shows that our way of seeing is deeply intertwined with how we interact with the world. We connect our perception to our experience in a physical world. AI models can't do that because they only exist in a computer."

AI can still learn from the human brain

The research thus touches on larger questions about the development of reliable and efficient AI. "As more sectors—from health care to robotics—use AI, it is becoming important that machines not only recognize what something is, but also understand what it can do," Groen explains. "For example, a robot that has to find its way in a disaster area, or a self-driving car that can tell apart a bike path from a driveway."

Groen also points out the sustainable aspect of AI. "Current AI training methods use a huge amount of energy and are often only accessible to large tech companies. More knowledge about how our brain works, and how the human brain processes certain information very quickly and efficiently, can help make AI smarter, more economical and more human-friendly."

More information: Clemens G. Bartnik et al, Representation of locomotive action affordances in human behavior, brains, and deep neural networks, *Proceedings of the National Academy of Sciences* (2025). DOI: 10.1073/pnas.2414005122
Journal information: Proceedings of the National Academy of Sciences
Provided by University of Amsterdam

JUNE 17, 2025

Wafer-scale accelerators could redefine AI

by David Danelski, University of California - Riverside

edited by Lisa Lock, reviewed by Robert Egan

The promise of a new type of computer chip that could reshape the future of artificial intelligence and be more environmentally friendly is explored in a technology review paper published by UC Riverside engineers in the journal *Device*.

Known as wafer-scale accelerators, these massive chips made by Cerebras are built on dinner plate-sized silicon wafers, in stark contrast to traditional graphics processing units, or GPUs, which are no bigger than a postage stamp.

The paper by a cross-disciplinary UCR team concludes that wafer-scale processors can deliver far more computing power with much greater energy efficiency—traits that are needed as AI models grow ever larger and more demanding.

"Wafer-scale technology represents a major leap forward," said Mihri Ozkan, a professor of electrical and computer engineering in UCR's Bourns College of Engineering and the paper's lead author. "It enables AI models with trillions of parameters to run faster and more efficiently than traditional systems."

In addition to Ozkan, co-authors include UCR graduate students Lily Pompa, Md Shaihan Bin Iqbal, Yiu Chan, Daniel Morales, Zixun Chen, Handing Wang, Lusha Gao, and Sandra Hernandez Gonzalez.

GPUs became essential tools for AI development because they can perform many computations at once—ideal for processing images, language, and data streams in parallel. The execution of thousands of parallel operations simultaneously allows for driverless cars to interpret the world around them to avoid collisions, for images to be generated from text, and for ChatGPT to suggest dozens of meal recipes from a specific list of ingredients.

But as AI model complexity increases, even high-end GPUs are starting to hit performance and energy limits.

"AI computing isn't just about speed anymore," Ozkan said. "It's about designing systems that can move massive amounts of data without overheating or consuming excessive electricity."

The UCR analysis compares today's standard GPU chips with wafer-scale systems like the Cerebras Wafer-Scale Engine 3 (WSE-3), which contains 4 trillion transistors and 900,000 AI-specific cores on a single wafer. Tesla's Dojo D1, another example, includes 1.25 trillion transistors and nearly 9,000 cores per module. These systems are engineered to eliminate the performance bottlenecks that occur when data must travel between multiple smaller chips.

"By keeping everything on one wafer, you avoid the delays and power losses from chip-to-chip communication," Ozkan said.

The paper also highlights technologies such as chip-on-wafer-on-substrate packaging, which could make wafer-scale designs more compact and easier to scale, with a potential 40-fold increase in computational density.

While these systems offer substantial advantages, they're not suited for every application. Wafer-scale processors are costly to manufacture and less flexible for smaller-scale tasks. Conventional GPUs, with their modularity and affordability, remain essential in many settings.

"Single-chip GPUs won't disappear," Ozkan said. "But wafer-scale accelerators are becoming indispensable for training the most advanced AI models."

The paper also addresses a growing concern in AI: sustainability. GPU-powered data centers use enormous amounts of electricity and water to stay cool. Wafer-scale processors, by reducing internal data traffic, consume far less energy per task.

For example, the Cerebras WSE-3 can perform up to 125 quadrillion operations per second while using a fraction of the power required by comparable GPU systems. Its architecture keeps data local, lowering energy draw and thermal output.

Meanwhile, NVIDIA's H100 GPU—the backbone of many modern data centers—offers flexibility and high throughput, but at greater energy cost. With an efficiency rate of about 7.9 trillion operations per second per watt, it also requires extensive cooling infrastructure, often involving large volumes of water.

"Think of GPUs as busy highways—effective, but traffic jams waste energy," Ozkan said. "Wafer-scale engines are more like monorails: direct, efficient, and less polluting."

Cerebras reports that inference workloads on its WSE-3 system use one-sixth the power of equivalent GPU-based cloud setups. The technology is already being used in climate simulations, sustainable engineering, and carbon-capture modeling.

"We're seeing wafer-scale systems accelerate sustainability research itself," Ozkan said. "That's a win for computing and a win for the planet."

However, heat remains a challenge. With thermal design power reaching 10,000 watts, wafer-scale chips require advanced cooling. Cerebras employs a glycol-based loop built into the chip package, while Tesla uses a coolant system that distributes liquid evenly across the chip surface.

The authors also emphasize that up to 86% of a system's total carbon footprint can come from manufacturing and supply chains, not just energy use. They advocate for recyclable materials and lower-emission alloys, along with full lifecycle design practices.

"Efficiency starts at the factory," Ozkan said. "To truly lower computing's impact, we need to rethink the whole process—from wafer to waste. This review is the result of a deep interdisciplinary collaboration. We hope it serves as a roadmap for researchers, engineers, and policymakers navigating the future of AI hardware."

More information: Mihrimah Ozkan et al, Performance, efficiency, and cost analysis of wafer-scale AI accelerators vs. single-chip GPUs, *Device* (2025). DOI: 10.1016/j.device.2025.100834
Journal information: Device
Provided by University of California - Riverside

JUNE 17, 2025

Lost in the middle: How LLM architecture and training data shape AI's position bias

by Adam Zewe, Massachusetts Institute of Technology

edited by Stephanie Baum, reviewed by Andrew Zinin

Three types of attention masks and their corresponding directed graphs G used in the analysis (self-loops are omitted for clarity). A directed edge from token j to i indicates that i attends to j. The center node(s) (Definition 3.1), highlighted in yellow, represent tokens that can be directly or indirectly attended to by all other tokens in the sequence. As depicted in the top row, the graph-theoretic formulation captures both direct and indirect contributions of tokens to the overall context, providing a comprehensive view of the token interactions under multi-layer attention. Credit: *arXiv* (2025). DOI: 10.48550/arxiv.2502.01951
Research has shown that large language models (LLMs) tend to overemphasize information at the beginning and end of a document or conversation, while neglecting the middle.

This "position bias" means that if a lawyer is using an LLM-powered virtual assistant to retrieve a certain phrase in a 30-page affidavit, the LLM is more likely to find the right text if it is on the initial or final pages.

MIT researchers have discovered the mechanism behind this phenomenon.

They created a theoretical framework to study how information flows through the machine-learning architecture that forms the backbone of LLMs. They found that certain design choices which control how the model processes input data can cause position bias.

Their experiments revealed that model architectures, particularly those affecting how information is spread across input words within the model, can give rise to or intensify position bias, and that training data also contribute to the problem.

The work is published on the *arXiv* preprint server.

In addition to pinpointing the origins of position bias, their framework can be used to diagnose and correct it in future model designs.

This could lead to more reliable chatbots that stay on topic during long conversations, medical AI systems that reason more fairly when handling a trove of patient data, and code assistants that pay closer attention to all parts of a program.

"These models are black boxes, so as an LLM user, you probably don't know that position bias can cause your model to be inconsistent. You just feed it your documents in whatever order you want and expect it to work. But by understanding the underlying mechanism of these black-box models better, we can improve them by addressing these limitations," says Xinyi Wu, a graduate student in the MIT Institute for Data, Systems, and Society (IDSS) and the Laboratory for Information and Decision Systems (LIDS), and first author of the paper.

Her co-authors include Yifei Wang, an MIT postdoc; and senior authors Stefanie Jegelka, an associate professor of electrical engineering and computer science (EECS) and a member of IDSS and the Computer Science and Artificial Intelligence Laboratory (CSAIL); and Ali Jadbabaie, professor and head of the Department of Civil and Environmental Engineering, a core faculty member of IDSS, and a principal investigator in LIDS. The research will be presented at the International Conference on Machine Learning.

Analyzing attention

LLMs like Claude, Llama, and GPT-4 are powered by a type of neural network architecture known as a transformer. Transformers are designed to process sequential data, encoding a sentence into chunks called tokens and then learning the relationships between tokens to predict which words come next.

These models have gotten very good at this because of the attention mechanism, which uses interconnected layers of data processing nodes to make sense of context by allowing tokens to selectively focus on or attend to related tokens.

But if every token can attend to every other token in a 30-page document, that quickly becomes computationally intractable. So, when engineers build transformer models, they often employ attention-masking techniques that limit the words to which a token can attend. For instance, a causal mask only allows words to attend to those that came before it.

Engineers also use positional encodings to help the model understand the location of each word in a sentence, improving performance.

The MIT researchers built a graph-based theoretical framework to explore how these modeling choices, attention masks and positional encodings, could affect position bias.

"Everything is coupled and tangled within the attention mechanism, so it is very hard to study. Graphs are a flexible language to describe the dependent relationship among words within the attention mechanism and trace them across multiple layers," Wu says.

Their theoretical analysis suggested that causal masking gives the model an inherent bias toward the beginning of an input, even when that bias doesn't exist in the data.

If the earlier words are relatively unimportant for a sentence's meaning, causal masking can cause the transformer to pay more attention to its beginning anyway.

"While it is often true that earlier words and later words in a sentence are more important, if an LLM is used on a task that is not natural language generation, like ranking or information retrieval, these biases can be extremely harmful," Wu says.

As a model grows, with additional layers of attention mechanism, this bias is amplified because earlier parts of the input are used more frequently in the model's reasoning process.

They also found that using positional encodings to link words more strongly to nearby words can mitigate position bias. The technique refocuses the model's attention in the right place, but its effect can be diluted in models with more attention layers. These design choices are only one cause of position bias—some can come from training data the model uses to learn how to prioritize words in a sequence.

"If you know your data is biased in a certain way, then you should also finetune your model on top of adjusting your modeling choices," Wu says.

Lost in the middle

After they'd established a theoretical framework, the researchers performed experiments in which they systematically varied the position of the correct answer in text sequences for an information retrieval task.

The experiments showed a "lost-in-the-middle" phenomenon, where retrieval accuracy followed a U-shaped pattern. Models performed best if the right answer was located at the beginning of the sequence. Performance declined the closer it got to the middle before rebounding a bit if the correct answer was near the end.

Ultimately, their work suggests that using a different masking technique, removing extra layers from the attention mechanism, or strategically employing positional encodings could reduce position bias and improve a model's accuracy.

"By doing a combination of theory and experiments, we were able to look at the consequences of model design choices that weren't clear at the time. If you want to use a model in high-stakes applications, you must know when it will work, when it won't, and why," Jadbabaie says.

In the future, the researchers want to further explore the effects of positional encodings and study how position bias could be strategically exploited in certain applications.

"These researchers offer a rare theoretical lens into the attention mechanism at the heart of the transformer model. They provide a compelling analysis that clarifies longstanding quirks in transformer behavior, showing that attention mechanisms, especially with causal masks, inherently bias models toward the beginning of sequences. The paper achieves the best of both worlds—mathematical clarity paired with insights that reach into the guts of real-world systems," says Amin Saberi, professor and director of the Stanford University Center for Computational Market Design, who was not involved with this work.

More information: Xinyi Wu et al, On the Emergence of Position Bias in Transformers, *arXiv* (2025). DOI: 10.48550/arxiv.2502.01951
Journal information: arXiv
Provided by Massachusetts Institute of Technology

JUNE 17, 2025

From code to commands: Prompt training technique helps users speak AI's language

by Aaron Aupperlee, Carnegie Mellon University

edited by Sadie Harley, reviewed by Andrew Zinin

Credit: Pixabay/CC0 Public Domain
Today's generative artificial intelligence models can create everything from images to computer applications, but the quality of their output depends largely on the prompt a human user provides.

Carnegie Mellon University researchers have proposed a new approach for teaching everyday users how to create these prompts and improving their interactions with generative artificial intelligence models.

The method, called Requirement-Oriented Prompt Engineering (ROPE), shifts the focus of prompt writing from clever tricks and templates to clearly stating what the AI should do. As large language models (LLMs) improve, the importance of coding skills may wane while expertise in prompt engineering could rise.

"You need to be able to tell the model exactly what you want. You can't expect it to guess all your customized needs," said Christina Ma, a Ph.D. student in the Human-Computer Interaction Institute (HCII). "We need to train humans in prompt engineering skills. Most people still struggle to tell the AI exactly what they want. ROPE helps them do that."

Prompt engineering refers to the precise instructions — the prompts —a user gives a generative AI model to produce a desired output. The better a user is at prompt engineering, the more likely an AI model will produce what the user intended.

In "What Should We Engineer in Prompts? Training Humans in Requirement-Driven LLM Use," appearing in the *Association for Computing Machinery's Transactions on Computer-Human Interaction*, researchers describe their ROPE paradigm and a training module they created to teach and assess the method.

ROPE is a human-LLM partnering strategy where humans maintain agency and control of the goals by specifying requirements for LLM prompts. The paradigm focuses on the importance of crafting accurate and complete requirements to achieve better results, especially for complex, customized tasks.

To test ROPE, the researchers asked 30 people to write prompts for an AI model to complete two separate tasks as a pretest: create a tic-tac-toe game and design a tool to help people develop content outlines. Half of the participants then received training through ROPE and the rest watched a YouTube tutorial on prompt engineering. The groups then wrote prompts for a different game and a different chatbot as a posttest.

When researchers compared the results of the exercises, they found that participants who received the ROPE training outperformed the people who watched the YouTube tutorial. Scores from pretest to posttest rose by 20% for people who received the ROPE training and only 1% for those who did not.

"We not only proposed a new framework for teaching prompt engineering but also created a training tool to assess how well participants do and how well the paradigm works," said Ken Koedinger, a University Professor in the HCII. "It's not just our opinion that ROPE works. The training module backs it up."

Generative AI models have already altered the content of introductory programming and software engineering courses as traditional programming evolves into natural language programming. Instead of writing software, an engineer can write a prompt directing AI to develop the software.

This paradigm shift could create new opportunities for students, allowing them to work on more complex development tasks earlier in their studies and advancing the field.

The researchers did not design ROPE solely for software engineers. As humans continue to integrate AI into daily life, clearly communicating with machines will become an important aspect of digital literacy. Armed with knowing how to write successful prompts and an AI model up to the task, people without coding or software engineering backgrounds can create applications that will benefit them.

"We want to empower more end users from the general public to use LLMs to build chatbots and apps," Ma said. "If you have an idea, and you understand how to communicate the requirements, you can write a prompt that will create that idea."

The researchers have open-sourced their training tools and materials, aiming to make prompt engineering more accessible to nonexperts.

More information: Qianou Ma et al, What Should We Engineer in Prompts? Training Humans in Requirement-Driven LLM Use, *ACM Transactions on Computer-Human Interaction* (2025). DOI: 10.1145/3731756
Provided by Carnegie Mellon University

JUNE 17, 2025

Robots learn welding skills from humans to address welder shortage

by University of Nottingham edited by Lisa Lock, reviewed by Andrew Zinin

Credit: CC0 Public Domain

Robots could be the solution to filling the shortage of welders in the U.K., thanks to existing human expertise, a new study from the University of Nottingham has revealed.

The U.K.'s critical welder shortage threatens certain industries, from construction to aerospace, potentially impacting the economy and infrastructure. Declining vocational training and Brexit as it becomes more challenging to attract and retain skilled welders from the EU has further exacerbated this skills gap. According to Axiom Personnel, half of the nation's welders are due to retire by 2027.

In the new study, published in *Robotics and Computer-Integrated Manufacturing*, the authors asked whether robots could bridge this divide and if expert welder skills could be transferred to automated systems.

By developing a robotic welding system that learns from skilled welders, and building a skills library, the system could then tackle new, unseen welding tasks by intelligently combining learned skills, demonstrated successfully with both expert and novice welders.

In the proposed approach, proficient welders execute basic tasks, such as welding simple lines or arcs, while their actions are recorded using an operation tracking system. Then key welding parameters, such as torch traveling speed, welding arc length, welding angle, welding current, and wire feeding rate, are extracted and stored in a skill library.

Experiments have also been conducted to verify the system. A skilled welder was asked to weld linear and arc-shaped grooves on stainless steel workpieces, while the welder's skills were tracked, extracted, and stored digitally.

These skills were further used to plan the robotic welding system to execute new complex tasks, such as polynomial curves—curves with a lot of bends. Welding results from the robot show a quality that is on par with that of a skilled welder, effectively saving time and resources in the long term.

"Our aim is not to replace expert welders with robots. Instead, we seek to enable robots to conduct repetitive or hazardous tasks after learning skills from expert welders to increase the productivity, thereby allowing the experts to focus on more creative tasks. The proposed methodology in this research is expandable to other manual operations such as assembly, polishing. Hence it can benefit several industrial sectors," says Abdelkhalick Mohammad.

Robotic welding systems are pivotal in various manufacturing sectors, such as aerospace, construction, automotive, and maritime industries, due to their ability to operate in challenging environments with fewer physical constraints compared to human welders.

However, their lack of process knowledge and adaptability leads to a heavy reliance on experienced technicians for process planning. To reduce these challenges, this new robotic welding system is proposed, focusing on learning from manual operations.

More information: Junfu Zhou et al, Teaching robots to weld by leveraging human expertise, *Robotics and Computer-Integrated Manufacturing* (2025). DOI: 10.1016/j.rcim.2025.103027
Provided by University of Nottingham

JUNE 17, 2025

AI 'reanimations': Making facsimiles of the dead raises ethical quandaries

by Nir Eisikovits, Daniel J. Feldman, The Conversation

edited by Lisa Lock, reviewed by Robert Egan

This screenshot of an AI-generated video depicts Christopher Pelkey, who was killed in 2021. Credit: Stacey Wales/YouTube
Christopher Pelkey was shot and killed in a road range incident in 2021. On May 8, 2025, at the sentencing hearing for his killer, an AI video reconstruction of Pelkey delivered a victim impact statement. The trial judge reported being deeply moved by this performance and issued the maximum sentence for manslaughter.

As part of the ceremonies to mark Israel's 77th year of independence on April 30, 2025, officials had planned to host a concert featuring four iconic Israeli singers. All four had died years earlier. The plan was to conjure them using AI-generated sound and video. The dead performers were supposed to sing alongside Yardena Arazi, a famous and still very much alive artist. In the end, Arazi pulled out, citing the political atmosphere, and the event didn't happen.

In April, the BBC created a deep-fake version of the famous mystery writer Agatha Christie to teach a "maestro course on writing." Fake Agatha would instruct aspiring murder mystery authors and "inspire" their "writing journey."

The use of artificial intelligence to "reanimate" the dead for a variety of purposes is quickly gaining traction. Over the past few years, we've been studying the moral implications of AI at the Center for Applied Ethics at the University of Massachusetts, Boston, and we find these AI reanimations to be morally problematic.

Before we address the moral challenges the technology raises, it's important to distinguish AI reanimations, or deepfakes, from so-called griefbots. Griefbots are chatbots trained on large swaths of data the dead leave behind—social media posts, texts, emails, videos. These chatbots mimic how the departed used to communicate and are meant to make life easier for surviving relations. The deepfakes we are discussing here have other aims; they are meant to promote legal, political and educational causes.

Moral quandaries

The first moral quandary the technology raises has to do with consent: Would the deceased have agreed to do what their likeness is doing? Would the dead Israeli singers have wanted to sing at an independence ceremony organized by the nation's current government? Would Pelkey, the road-rage victim, be comfortable with the script his family wrote for his avatar to recite? What would Christie think about her AI double teaching that class?

The answers to these questions can only be deduced circumstantially—from examining the kinds of things the dead did and the views they expressed when alive. And one could ask if the answers even matter. If those in charge of the estates agree to the reanimations, isn't the question settled? After all, such trustees are the legal representatives of the departed.

But putting aside the question of consent, a more fundamental question remains.

What do these reanimations do to the legacy and reputation of the dead? Doesn't their reputation depend, to some extent, on the scarcity of appearance, on the fact that the dead can't show up anymore? Dying can have a salutary effect on the reputation of prominent people; it was good for John F. Kennedy, and it was good for Israeli Prime Minister Yitzhak Rabin.

The fifth-century B.C. Athenian leader Pericles understood this well. In his famous Funeral Oration, delivered at the end of the first year of the Peloponnesian War, he asserts that a noble death can elevate one's reputation and wash away their petty misdeeds. That is because the dead are beyond reach and their mystique grows postmortem. "Even extreme virtue will scarcely win you a reputation equal to" that of the dead, he insists.

Do AI reanimations devalue the currency of the dead by forcing them to keep popping up? Do they cheapen and destabilize their reputation by having them comment on events that happened long after their demise?

In addition, these AI representations can be a powerful tool to influence audiences for political or legal purposes. Bringing back a popular dead singer to legitimize a political event and reanimating a dead victim to offer testimony are acts intended to sway an audience's judgment.

It's one thing to channel a Churchill or a Roosevelt during a political speech by quoting them or even trying to sound like them. It's another thing to have "them" speak alongside you. The potential of harnessing nostalgia is supercharged by this technology. Imagine, for example, what the Soviets, who literally worshiped Lenin's dead body, would have done with a deepfake of their old icon.

Good intentions

You could argue that because these reanimations are uniquely engaging, they can be used for virtuous purposes. Consider a reanimated Martin Luther King Jr., speaking to our currently polarized and divided nation, urging moderation and unity. Wouldn't that be grand? Or what about a reanimated Mordechai Anielewicz, the commander of the Warsaw Ghetto uprising, speaking at the trial of a Holocaust denier like David Irving?

But do we know what MLK would have thought about our current political divisions? Do we know what Anielewicz would have thought about restrictions on pernicious speech? Does bravely campaigning for civil rights mean we should call upon the digital ghost of King to comment on the impact of populism? Does fearlessly fighting the Nazis mean we should dredge up the AI shadow of an old hero to comment on free speech in the digital age?

Even if the political projects these AI avatars served were consistent with the deceased's views, the problem of manipulation—of using the psychological power of deepfakes to appeal to emotions—remains.

But what about enlisting AI Agatha Christie to teach a writing class? Deepfakes may indeed have salutary uses in educational settings. The likeness of Christie could make students more enthusiastic about writing. Fake Aristotle could improve the chances that students engage with his austere Nicomachean Ethics. AI Einstein could help those who want to study physics get their heads around general relativity.

But producing these fakes comes with a great deal of responsibility. After all, given how engaging they can be, it's possible that the interactions with these representations will be all that students pay attention to, rather than serving as a gateway to exploring the subject further.

Living on in the living

In a poem written in memory of W.B. Yeats, W.H. Auden tells us that after the poet's death, Yeats "became his admirers." His memory was now "scattered among a hundred cities," and his work subject to endless interpretation: "The words of a dead man are modified in the guts of the living."

The dead live on in the many ways we reinterpret their words and works. Auden did that to Yeats, and we're doing it to Auden right here. That's how people stay in touch with those who are gone. In the end, we believe that using technological prowess to concretely bring them back disrespects them, and perhaps more importantly, is an act of disrespect to ourselves—to our capacity to abstract, think and imagine.

JUNE 18, 2025

Seeing through a new LENS allows brain-like navigation in robots

by Queensland University of Technology edited by Sadie Harley, reviewed by Robert Egan

Dr. Adam Hines, with his 'green' robot. l/r- Dr. Tobias Fischer, Dr. Adam Hines and Professor Michael Milford. Credit: QUT
QUT robotics researchers have developed a new robot navigation system that mimics the neural processes of the human brain and uses less than 10% of the energy required by traditional systems.

In a study published in the journal *Science Robotics*, the researchers detail a new system which they call LENS—Locational Encoding with Neuromorphic Systems. The paper is titled "A compact neuromorphic system for ultra energy-efficient, on-device robot localization."

LENS uses brain-inspired computing to set a new, low-energy benchmark for robotic place recognition.

The research, conducted by first author neuroscientist Dr. Adam Hines along with Professor Michael Milford and Dr. Tobias Fischer, all from the QUT Center of Robotics and the QUT School of Electrical Engineering and Robotics, uses a system called neuromorphic computing

"To run these neuromorphic systems, we designed specialized algorithms that learn more like humans do, processing information in the form of electrical spikes, similar to the signals used by real neurons," Dr. Hines said.

"Energy constraints are a major challenge in real-world robotics, especially in fields like search and rescue, space exploration and underwater navigation.

"By using neuromorphic computing, our system reduces the energy requirements of visual localization by up to 99%, allowing robots to operate longer and cover greater distances on limited power supplies.

"We have known neuromorphic systems could be more efficient, but they're often too complex and hard to use in the real world—we developed a new system that we think will change how they are used with robots."

In the study, the researchers developed LENS, a system that was able to recognize locations along an 8km journey but using only 180KB of storage—almost 300 times less than other systems.

LENS combines a brain-like spiking neural network with a special camera that only reacts to movement and a low-power chip, all on one small robot.

"This system demonstrates how neuromorphic computing can achieve real-time, energy-efficient location tracking on robots, opening up new possibilities for low-power navigation technology," Dr. Hines said.

"Lower energy consumption can allow remotely operated robots to explore for longer and further.

"Our system enables robots to localize themselves using only visual information, in a way that is both fast and energy efficient."

Dr. Fischer, ARC DECRA Fellow, said the key innovation in the LENS system was a new algorithm that exploited two types of promising bio-inspired hardware: sensing, via a special type of camera known as an "event camera," and computing, via a neuromorphic chip.

"Rather than capturing a full image of the scene that takes in every detail in each frame, an event camera continuously senses changes and movement every microsecond," Dr. Fischer said.

"The camera detects changes in brightness at each pixel, closely replicating how our eyes and brain process visual information.

"Knowing where you are, also known as visual place recognition, is essential for both humans and robots.

"While people use visual cues effortlessly, it's a challenging task for machines."

Professor Michael Milford, director of the QUT Center for Robotics, said the study was representative of a key theme of research conducted by the center's researchers.

"Impactful robotics and tech means both pioneering ground-breaking research, but also doing all the translational work to ensure it meets end user expectations and requirements," Professor Milford said.

"You can't just do one or the other.

"This study is a great example of working towards energy-efficient robotic systems that provide end-users with the performance and endurance they require for those robots to be useful in their application domains."

More information: A compact neuromorphic system for ultra energy-efficient, on-device robot localization, *Science Robotics* (2025). DOI: 10.1126/scirobotics.ads3968
Journal information: Science Robotics
Provided by Queensland University of Technology

JUNE 19, 2025

Justice at stake as generative AI enters the courtroom

by Thomas URBAINedited by Andrew Zinin

Generative artificial intelligence has been used in the US legal system by judges performing research, lawyers filing appeals and parties involved in cases who wanted help expressing themselves in court.
Generative artificial intelligence (GenAI) is making its way into courts despite early stumbles, raising questions about how it will influence the legal system and justice itself.

Judges use the technology for research, lawyers utilize it for appeals and parties involved in cases have relied on GenAI to help express themselves in court.

"It's probably used more than people expect," said Daniel Linna, a professor at the Northwestern Pritzker School of Law, about GenAI in the US legal system.

"Judges don't necessarily raise their hand and talk about this to a whole room of judges, but I have people who come to me afterward and say they are experimenting with it."

In one prominent instance, GenAI enabled murder victim Chris Pelkey to address an Arizona courtroom—in the form of a video avatar—at the sentencing of the man convicted of shooting him dead in 2021 during a clash between motorists.

"I believe in forgiveness," said a digital proxy of Pelkey created by his sister, Stacey Wales.

The judge voiced appreciation for the avatar, saying it seemed authentic.

"I knew it would be powerful," Wales told AFP, "that that it would humanize Chris in the eyes of the judge."

The AI testimony, a first of its kind, ended the sentencing hearing at which Wales and other members of the slain man's family spoke about the impact of the loss.

Since the hearing, examples of GenAI being used in US legal cases have multiplied.

"It is a helpful tool and it is time-saving, as long as the accuracy is confirmed," said attorney Stephen Schwartz, who practices in the northeastern state of Maine.

"Overall, it's a positive development in jurisprudence."

Schwartz described using ChatGPT as well as GenAI legal assistants, such as LexisNexis Protege and CoCounsel from Thomson Reuters, for researching case law and other tasks.

"You can't completely rely on it," Schwartz cautioned, recommending that cases proffered by GenAI be read to ensure accuracy.

"We are all aware of a horror story where AI comes up with mixed-up case things."

The technology has been the culprit behind false legal citations, far-fetched case precedents, and flat-out fabrications.

In early May, a federal judge in Los Angeles imposed $31,100 in fines and damages on two law firms for an error-riddled petition drafted with the help of GenAI, blasting it as a "collective debacle."

The tech is also being relied on by some who skip lawyers and represent themselves in court, often causing legal errors.

And as GenAI makes it easier and cheaper to draft legal complaints, courts already overburdened by caseloads could see them climb higher, said Shay Cleary of the National Center for State Courts.

"Courts need to be prepared to handle that," Cleary said.

Transformation

Law professor Linna sees the potential for GenAI to be part of the solution though, giving more people the ability to seek justice in courts made more efficient.

"We have a huge number of people who don't have access to legal services," Linna said.

"These tools can be transformative; of course we need to be thoughtful about how we integrate them."

Federal judges in the US capitol have written decisions noting their use of ChatGPT in laying out their opinions.

"Judges need to be technologically up-to-date and trained in AI," Linna said.

GenAI assistants already have the potential to influence the outcome of cases the same way a human law clerk might, reasoned the professor.

Facts or case law pointed out by GenAI might sway a judge's decision, and could be different than what a legal clerk would have come up with.

But if GenAI lives up to its potential and excels at finding the best information for judges to consider, that could make for well-grounded rulings less likely to be overturned on appeal, according to Linna.

JUNE 19, 2025

Hyper-realistic AI technology creates avatars from a single photo

by National Research Council of Science and Technology edited by Gaby Clark, reviewed by Robert Egan

AI Avatar. Credit: Electronics and Telecommunications Research Institute(ETRI)
Electronics and Telecommunications Research Institute (ETRI) has developed hyper-realistic AI technology that can create an avatar that speaks naturally like a real person using only a single portrait photo.

The technology is being seen as a next-generation interface that enables intuitive interaction between vehicles and humans in preparation for the era of fully autonomous driving, and is expected to spread across the digital human industry.

While traditional current speech-driven AI assistants in office environments or navigation systems in vehicles are limited to simply carrying out commands, ETRI's hyper-realistic AI avatars have sophisticated facial expressions and mouth movements that enable natural, human-like conversations.

This allows for a more human-centered human-machine interaction, such as an in-vehicle AI driver talking to the driver or interacting with pedestrians.

The core of this technology is a unique algorithm that, unlike traditional generative AI, selectively learns and synthesizes parts of the face that are directly related to utterance, such as the lips and chin. This approach reduces unnecessary information learning and allows for more sophisticated facial expressions, including mouth shapes, teeth, and skin wrinkles.

ETRI explained that the technology has demonstrated superior performance in terms of synthetic visual quality and lip synchronization accuracy as presented at major international conferences such as CVPR and AAAI.

In addition to autonomous vehicles, this technology can be utilized in various industries such as kiosks, bank counters, news presentations, advertising models, and is expected to drive innovation in the AI-based digital human industry.

ETRI's Mobility User Experience Research Section is currently focusing on human-machine interaction (HMI) technologies, and is also developing AI-based driver interface technologies that analyze driver and pedestrian emotions, fatigue, concentration, etc.

Daesub Yoon, Director of the Mobility User Experience Research Section, said, "As mobility technology becomes more advanced, the elderly and socially disadvantaged may be marginalized. We hope that this AI avatar technology will contribute to improving digital literacy and make smart mobility services more accessible to all."

And Senior Researcher Daewoong Choi also said, "We plan to further advance our generative AI technology so that AI avatars can naturally talk and move like real people. In the future, we're aiming for interactions that can replace some human labor for ordering, consulting, and more."

The technology is currently registered on the ETRI Technology Transfer site as "A Framework for Photorealistic Talking Face Generation." The researchers will also actively pursue technology transfer and strategies for commercialization in various industries.

Provided by National Research Council of Science and Technology

---

JUNE 19, 2025

Researchers are teaching AI to see more like humans

by National Science Foundation edited by Sadie Harley, reviewed by Robert Egan

At Brown University, an innovative new project is revealing that teaching artificial intelligence to perceive things more like people may begin with something as simple as a game. The project invites participants to play an online game called Click Me, which helps AI models learn how people see and interpret images. While the game is fun and accessible, its purpose is more ambitious: to understand the root causes of AI errors and to systematically improve how AI systems represent the visual world.

Over the past decade, AI systems have become more powerful and widely used, particularly in tasks like recognizing images. For example, these systems can identify animals, objects or diagnose medical conditions from images. However, they sometimes make mistakes that humans rarely do.

For instance, an AI algorithm might confidently label a photo of a dog wearing sunglasses as a completely different animal or fail to recognize a stop sign if it's partially covered by graffiti. As these models become larger and more complex, these kinds of errors become more frequent, revealing a growing gap between how AI and humans perceive the world.

Recognizing this challenge, researchers propose to combine insights from psychology and neuroscience with machine learning to create the next generation of human-aligned AI. Their goal is to understand how people process visual information and translate those patterns into algorithms that guide AI systems to act in similar ways.

The Click Me game plays a central role in this vision. In the game, participants click on parts of an image they believe will be most informative for the AI to recognize. The AI only sees the parts of the image that have been clicked. Therefore, players are encouraged to think strategically about the most informative parts of the image rather than clicking at random to maximize the AI's learning.

The AI-human alignment occurs at a later stage, during which the AI is trained to categorize images. In this "neural harmonization" procedure, the researchers force the AI to focus on the same image features that humans had identified — those clicked during the game — to make sure its visual recognition strategy aligns with that of humans.

What makes this project especially remarkable is how successfully it has engaged the public. The team has attracted thousands of people to participate in Click Me, helping it gain attention across platforms like Reddit and Instagram, and generating tens of millions of interactions with the website to help train the AI model. This type of large-scale public participation allows the research team to rapidly collect data on how people perceive and evaluate visual information.

At the same time, the team has also developed a new computational framework to train AI models using this kind of behavioral data. By aligning AI response times and choices with those of humans, the researchers can build systems that not only match what humans decide, but also how long they take to decide. This leads to a more natural and interpretable decision-making process.

The practical applications of this work are wide-ranging. In medicine, for instance, doctors need to understand and trust the AI tools that assist with diagnoses. If AI systems can explain their conclusions in ways that match human reasoning, they become more reliable and easier to integrate into care.

Similarly, in self-driving cars, AI that better understands how humans make visual decisions can help predict driver behavior and prevent accidents. Beyond these examples, human-aligned AI could improve accessibility tools, educational software and decision support across many industries. Importantly, this work also sheds light on how the human brain works.

By emulating human vision in AI systems, the researchers have been able to develop more accurate models of human visual perception than were previously available.

This initiative underscores why federal support for foundational research matters. Through NSF's investment, researchers are advancing the science of AI and its relevance to society. The research not only pushes the boundaries of knowledge but also delivers practical tools that can improve the safety and reliability of the technologies we use daily.

Provided by National Science Foundation

JUNE 20, 2025

All-topographic neural networks more closely mimic the human visual system

by Ingrid Fadelli, Phys.org edited by Robert Egan

All-TNNs better approximate spatial biases in human visual behavior. Credit: *Nature Human Behaviour* (2025). DOI: 10.1038/s41562-025-02220-7
Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are designed to partly emulate the functioning and structure of biological neural networks. As a result, in addition to tackling various real-world computational problems, they could help neuroscientists and psychologists to better understand the underpinnings of specific sensory or cognitive processes.

Researchers at Osnabrück University, Freie Universität Berlin and other institutes recently developed a new class of artificial neural networks (ANNs) that could mimic the human visual system better than CNNs and other existing deep learning algorithms. Their newly proposed, visual system-inspired computational techniques, dubbed all-topographic neural networks (All-TNNs), are introduced in a paper published in *Nature Human Behaviour*.

"Previously, the most powerful models for understanding how the brain processes visual information were derived off of AI vision models," Dr. Tim Kietzmann, senior author of the paper, told Tech Xplore.

"These are often convolutional in nature—a machine learning hack that allows the corresponding neural networks to search for the exact same feature everywhere in the visual input. This approach is very powerful: what you learn in one location of space can be transferred to all others. However, this is something the brain cannot do (the brain cannot 'copy' and 'paste' information from one location of the cortex to another)."

In addition to performing some actions that the primate brain is incapable of performing, CNNs also organize information differently from biological neural networks. In contrast with CNNs, the brain is retinotopically organized, which means that visual signals travel from the retina to the visual cortex (a region of the brain's outer layer known to process visual information).

"The brain also exhibits a systematic relationship between the types of features it is responding to and the location at which it is searching for them," said Kietzmann.

"This interrelation of space and feature along the cortical surface is an essential aspect of visual processing, but, as stated above, this feature is not considered in machine learning. To solve this shortcoming, we developed a biologically more realistic model class 'All topographic neural networks,' in which feature selectivity is spatially organized across a 'cortical sheet,' i.e., a 2D surface in which neighboring features are bound to be similar, but vary across larger distances)."

Most computational approaches commonly used to model how the human visual system processes natural images rely on deep neural networks (DNNs), such as CNNs. These are powerful models that can be trained to classify visual data, such as brain imaging scans, or to identify specific objects in images.

"The problem with these models is that they are often quite far removed from biology, and newer ML models, despite being more powerful, also stopped being better models of visual processing in the brain (a relationship that held true in the past)," explained Kietzmann.

"Across a series of papers, my lab demonstrates ways in which we can change the ML models to be better models of biology. For example, by training on better image datasets, by including recurrent connectivity in the network architecture, by considering what task the models should be trained for, and most recently, by considering that the brain has feature detectors aligned across the cortical surface."

Kietzmann and his colleagues demonstrated that the new models they developed, based on (All-TNNs), mirror the human visual system more closely than CNNs and other DNNs. This is because they do not only replicate the principles underpinning the organization of the visual cortex, but they also capture human behavioral patterns better than previously developed models.

In the future, All-TNNs could be used to carry out neuroscience and psychology studies, potentially shedding new light on the neural underpinnings of the human visual system. For instance, they could help to better understand how the arrangement of feature selectivity across the cortex, also known as topography, influences human perception and behavior.

"We are currently trying to improve training to be more efficient in terms of task performance, as topographic networks are parameter rich, compared to their convolutional counterparts," added Kietzmann.

"In addition, we currently need to steer the models towards smooth feature selectivity across space—a key feature of cortical topography. However, biology has likely developed implicit mechanisms that make the cortical selectivity smooth. Finding out which aspects allow for this to happen is a main area of research that we hope to be able to contribute to."

Written for you by our author Ingrid Fadelli, edited by Robert Egan —this article is the result of careful human work. We rely on readers like you to keep independent science journalism alive. If this reporting matters to you, please consider a donation (especially monthly). You'll get an **ad-free** account as a thank-you.

More information: Zejin Lu et al, End-to-end topographic networks as models of cortical map formation and human visual behaviour, *Nature Human Behaviour* (2025). DOI: 10.1038/s41562-025-02220-7.
Journal information: Nature Human Behaviour

JUNE 23, 2025

The law relies on being precise. AI is disrupting that

by Andrew Lim and Jeannie Marie Paterson, University of Melbourne

edited by Lisa Lock, reviewed by Andrew Zinin

The use of AI is spreading throughout all professions and industries—and law is no exception.

Lawyers have been innovative in their use of large language models (LLMs), but we're also seeing examples of them being caught out by the language capabilities of generative AI.

Don't rely on AI

Refined AI models, combined with retrieval or "RAG" systems, have the capacity to summarize, review and analyze documents.

But increasingly, courts and legal regulators are warning lawyers not to rely on generative AI without checking their work—and in some instances are advising them not to use it at all.

In fact, we've already seen courts admonish lawyers for submitting legal documents that include content fabricated by AI.

And in Australia, courts themselves have been cautious about the use of the technology.

Ordinary meaning and generative AI

Some courts around the world are experimenting with using generative AI.

In England, Lord Justice Birss advised that he used ChatGPT to summarize an area of the law. In the U.S. case of *Snell v United Specialty Insurance Co*, Justice Kevin Newsom used ChatGPT to determine the plain and ordinary meaning of a contentious term in an insurance policy.

When interpreting documents to decide their legal meaning, courts typically look to the ordinary meaning of words, as well as the context in which they are used.

But how do judges determine the ordinary meaning of a word?

One approach is to ask "ordinary people."

For most of the 19th century, the U.S. Supreme Court mandated that its justices ride town-to-town hearing cases to give them exposure to everyday citizens and conditions outside the capital.

Judges often consult a dictionary. Or perhaps they ask ChatGPT.

One tool among many

Let's go back to the Snell case in the U.S. mentioned earlier. In mid-2024, Justice Kevin Newsom was tasked with deciding whether an insurance policy's coverage of "landscaping" included the installation of a trampoline.

Justice Newsom checked three dictionaries and found three very different answers.

His Honor considered a "visceral, gut-instinct" feeling, only to decide it didn't seem very legally compelling. Instead, His Honor asked ChatGPT.

The possibility of generative AI providing an "ordinary meaning" is in some ways compelling.

After all, these models are trained on vast corpuses of the English language—books, newspapers, user prompts—covering all sorts of speech in all sorts of contexts.

Their reading is not limited by background, interests or age. In this sense, ChatGPT might be seen as representing an amalgam of the ordinary person in a way that judges, or even the compilers of dictionaries, are not.

All that said, the judge ended by sounding a note of caution. In his Honor's view, LLMs should only be "one tool among many," held up and tested against historical context, common sense and dictionary meanings.

Not an oracle of ordinary meaning

There are also important differences between generative AI and dictionaries, which limit their role as oracles of ordinary meaning.

As numerous lawyers have discovered, not only does generative AI hallucinate, it can also be sycophantic, providing the answer that, as suggested by the context, is what the person providing the prompt wants to hear.

This means the tools are likely to offer up confident-but-rubbish answers by attempting to mimic human speech.

Another concern is around transparency.

We know how dictionaries are compiled and the process is scrupulously documented. We have no clear understanding of the training data used in the free general-purpose AI models.

And, unlike dictionaries, which use a collaborative process for compiling content, much of the AI training data is often used without the knowledge or consent of the original author.

From common meaning to judgment

Courts, businesses, governments and firms are continuing to experiment with generative AI. Tech firms and developers are continuing to find ways to refine and improve their outputs.

For lawyers and courts, accuracy matters. But so do other values, like transparency, fairness and accountability.

These are part of Australia's ethical AI framework and central to the administration of justice.

A judge's core responsibility in interpreting texts has remained unchanged across the centuries: resolving the tension between clinical legal language and the often-messy nature of our lived realities.

The allure of generative AI's data-backed outputs should not distract from the fact that these decisions invariably rely on judgment and context.

While it may present new pieces of information in making these decisions, generative AI cannot be treated as any more authoritative or reliable than any other source—and certainly no more ethically compelling.

Provided by University of Melbourne

JUNE 23, 2025

AI applications are producing cleaner cities, smarter homes and more efficient transit

by Mohammadamin Ahmadfard, The Conversation edited by Lisa Lock, reviewed by Andrew Zinin

Credit: Pixabay/CC0 Public Domain

Artificial intelligence (AI) is quietly transforming how cities generate, store and distribute energy, acting as the invisible conductor that orchestrates cleaner, smarter and more resilient cities.

By integrating renewables—from solar panels and wind turbines to geothermal grids, hydrogen plants, electric vehicles and batteries—AI can enable cities to manage diverse energy sources as a single, intelligent system.

One striking example is the Oya Hybrid Power Station in South Africa. Here, AI-driven controls seamlessly coordinate solar, wind and battery storage to deliver reliable power to up to 320,000 households. Using AI makes this kind of integration not only possible, but dramatically more efficient.

Recent research shows AI can also optimize how batteries, solar and the grid interact in buildings. A 2023 study found that deep learning and real-time data helped a boarding school in Turin, Italy increase low-cost energy purchases and cut its electricity bill by more than half.

Cleaner, smarter energy grids

AI models are increasingly able to predict weather with greater precision. These predictions allow electric grid operators to plan hours ahead, storing excess energy in batteries or adjusting supply to meet demand before a storm or heat wave hits.

Using AI to respond strategically to weather is a game-changer. In Cambridge, England, a system called Aardvark uses satellite and sensor data to generate rapid, accurate forecasts of sun and wind patterns.

Unlike traditional supercomputer-driven weather models, Aardvark's AI can deliver precise local forecasts in minutes on an ordinary computer. This makes advanced weather prediction more accessible and affordable for cities, utilities and even smaller organizations—potentially transforming how communities everywhere plan for and respond to changing weather.

AI for smarter district heating and cooling

In Munich, Germany, AI is improving geothermal district heating by using underground sensors to monitor temperature and moisture levels in the ground.

The collected data feeds into a digital simulation model that helps optimize network operations. In more advanced versions, during winter cold snaps, such systems can suggest lowering flow to underused spaces like half-empty offices and boosting heat where demand is higher, such as in crowded apartments.

This intelligent, self-optimizing approach extends the life of equipment and delivers more warmth with the same energy input.

This is a breakthrough with enormous potential for cities in cold climates with established geothermal networks, such as Winnipeg in Canada and Iceland's Reykjavik.

Although these cities have not yet adopted AI-driven monitoring systems, they could benefit from AI's real-time improvements in efficiency, comfort and energy savings during harsh winters—a principle that holds true wherever geothermal district heating and cooling exists.

Smart buildings

Inside the home, AI-managed smart climate systems can factor in how many people are in each room, which appliances are in use, how much natural sunlight each space receives and how much electricity or heat a home's solar panels generate throughout the day.

Based on this, AI determines how to heat or cool rooms efficiently, and can transfer energy from one space to another, balancing comfort with minimal energy use.

Coastal cities and those in wind-heavy regions are using AI in other creative ways. In Orkney, Scotland, excess wind and tidal energy are converted into green hydrogen. Instead of letting that surplus power go to waste, an AI system called HyAI controls when to generate hydrogen based on wind forecasts, electricity prices and how full the hydrogen storage tanks are.

When winds are strong at night and electricity is cheap, the AI can divert surplus power to produce hydrogen and store it for later use. On calmer days, that stored hydrogen can power fuel cells or buses.

Energy storage

AI is transforming energy storage into a smart, revenue-generating force. In Finland, a startup called Capalo AI has developed Zeus VPP, an AI-powered virtual power plant that aggregates distributed batteries from homes, businesses and other sites.

Zeus VPP uses advanced forecasting and AI algorithms to decide when batteries should charge or discharge, factoring in energy prices, local consumption and weather forecasts. This enables battery owners to earn revenue by participating in electricity markets, while also supporting grid stability and making better use of renewable energy.

Utility companies are also using AI to monitor everything from high-voltage transmission lines to neighborhood transformers, dramatically increasing reliability.

AI-powered dynamic line rating adjusts how much electricity a line can carry in real time, boosting capacity by 15 to 30% when conditions allow. This helps utilities maximize the use of existing infrastructure instead of relying on costly upgrades.

At the local level, AI analyzes smart meter data to predict which transformers are overheating due to rising EV and heat pump use.

By forecasting these stress points, utilities can proactively upgrade equipment before failures happen—a shift from reactive to predictive maintenance that makes the grid stronger and cities more resilient.

AI-powered public transit and mobility

Transportation innovation is becoming part of the energy solution, with AI at the center of this transformation. In New York City, energy company Con Edison has installed major battery storage systems to help manage peak electricity demand and reduce reliance on polluting peaker plants, which supply energy only during high-demand periods.

More broadly, Con Edison is deploying advanced AI-powered analytics software across its electric grid—optimizing voltage, enhancing reliability and enabling predictive maintenance. Together, these efforts show how combining energy storage and AI-driven analytics can make even the world's busiest cities more resilient and efficient.

AI is also powering "vehicle-to-grid" innovations in California, where an AI-driven platform manages electric school buses that can supply stored energy back to the grid during periods of high demand.

By carefully managing when buses charge and discharge, these systems help keep the grid reliable and ensure vehicles are ready for their daily routes. As this technology expands, parked electric vehicles could serve as valuable backup resources for the electricity system.

AI for clean energy initiatives

AI is rapidly transforming cities by revolutionizing how energy is used and managed. Google, for example, has slashed cooling energy at its data centers by up to 40% using AI that fine-tunes fans, pumps and windows more efficiently than any human operator.

Organizations like the Electric Power Research Institute (EPRI), in collaboration with NVIDIA, Microsoft and others, have launched the Open Power AI Consortium, which is creating open-source AI tools for utilities worldwide.

These tools will enable even the most resource-constrained cities to deploy advanced AI capabilities, without having to start from scratch, helping to level the playing field and accelerate the global energy transition.

The result is not just cleaner air and lower energy bills, but a path to fewer blackouts and more resilient homes.

JUNE 23, 2025

AI is consuming more power than the grid can handle. Nuclear might be the answer

by Goran Calic, The Conversation edited by Lisa Lock, reviewed by Andrew Zinin

Credit: Pixabay/CC0 Public Domain

New partnerships are forming between tech companies and power operators—ones that could reshape decades of misconceptions about nuclear energy.

Last year, Meta (Facebook's parent company) put out a call for nuclear proposals, Google agreed to buy new nuclear reactors from Kairos Power, Amazon partnered with Energy Northwest and Dominion Energy to develop nuclear energy and Microsoft committed to a 20-year deal to restart Unit 1 of the Three Mile Island nuclear plant.

At the center of these partnerships is artificial intelligence's voracious appetite for electricity. One Google search uses about as much electricity as turning on a household light for 17 seconds. Asking a Generative AI model like ChatGPT a single question is equivalent to leaving that light on for 20 minutes.

Having GenAI generate an image can draw about 6,250 times more electricity, roughly the energy of fully charging a smartphone, or enough to keep the same light bulb on for 87 consecutive days.

The hundreds of millions of people now using AI have effectively added the equivalent of millions of new homes to the power grid. And demand is only growing. The challenge for tech companies is that few sources of electricity are well-suited to AI.

The grid wasn't ready for AI

AI requires vast amounts of computational power running around the clock, often housed in energy-intensive data centers.

Renewable energy sources such as solar and wind provide intermittent energy, meaning they don't guarantee the constant power supply these data centers require. These centers must be online 24/7, even when the sun isn't shining and the wind isn't blowing.

Fossil fuels can run continuously, but they carry their own risks. They have significant environmental impacts. Fuel prices can be unpredictable, as exemplified by the gas price spikes due to the war in Ukraine, and the long-term availability of fossil fuels is uncertain.

Major tech companies like Google, Amazon and Microsoft say they are committed to eliminating CO_2 emissions, making fossil fuels a poor long-term fit for them.

This has pushed nuclear energy back into the conversation. Nuclear energy is a good fit because it provides electricity around the clock, maximizing the use of expensive data centers. It's also clean, allowing tech companies to meet their low CO_2 commitments. Lastly, nuclear energy has very low fuel costs, which allows tech companies to plan their costs far into the future.

However, nuclear energy has its own set of problems that have historically been hard to solve—problems that tech companies may now be uniquely positioned to overcome.

Is nuclear energy making a comeback?

Nuclear power has long been considered too costly and too slow to build. The estimated cost of a 1.1 gigawatt nuclear power facility is about US$7.77 billion, but can run higher. The recently completed Vogtle Units 3 and 4 in the state of Georgia, for example, cost US$36.8 billion combined.

Historically, nuclear energy projects have been hard to justify because of their high upfront costs. Like solar and wind power, nuclear energy has relatively low operating costs once a plant is up and running. The key difference is scale: unlike solar panels, which can be installed on individual rooftops, the kind of nuclear reactors tech companies require can't be built small.

Yet this cost is now more palatable when compared to the expense of AI data centers, which are both more costly and entirely useless without electricity. The first phase of OpenAI and SoftBank's Stargate AI project will cost US$100 billion and could be entirely powered by a single nuclear plant.

Nuclear power plants also take a long time to build. A 1.1 gigawatt reactor takes, on average, 7.5 years in the U.S. and 6.3 years globally. Projects with such long timelines require confidence in long-term electricity demand, something traditional utilities struggle to predict.

To solve the problem of long-range forecasting, tech companies are incentivizing power providers by guaranteeing they'll purchase electricity far into the future.

These companies are also literally and financially moving closer to nuclear power, either by acquiring nuclear energy companies or locating their data centers next to nuclear power plants.

Destigmatizing nuclear energy

One of the biggest challenges facing nuclear energy is the perception that it's dangerous and dirty. Per gigawatt-hour of electricity, nuclear produces only six metric tons of CO_2. In comparison, coal produces 970, natural gas 720 and hydropower 24. Nuclear even has lower emissions than wind and solar, which produce 11 and 53 metric tons of CO_2, respectively.

Nuclear energy is also among the safest energy sources. Per gigawatt-hour, it causes 820 times fewer deaths than coal, 43 times fewer than hydropower and roughly the same as wind and solar.

Still, nuclear energy remains stigmatized, largely because of persistent misconceptions and outdated beliefs about nuclear waste and disasters. For instance, while many public concerns remain about nuclear waste, existing storage solutions have been used safely for decades and are supported by a strong track record and scientific consensus.

Similarly, while the Fukushima disaster in Japan displaced thousands of people and was extremely costly (total costs of the disaster are expected at about US$188 billion), not a single person died of radiation exposure after the accident, a United Nations Scientific Committee of 80 international experts found.

For decades, there was little effort to correct public perceptions about nuclear fears because it wasn't seen as necessary or profitable. Coal, gas and renewables were sufficient to meet the demand required of them. But that's now changing.

With AI's energy needs soaring, Big Tech has classified nuclear energy as green and the World Bank has agreed to lift its longstanding ban on financing nuclear projects.

Big Tech's billion-dollar bet on nuclear

The world has long lived with two nuclear dilemmas. The first is that, despite being one of the safest and cleanest forms of energy, nuclear was perceived as one of the most dangerous and dirtiest.

The second is that upgrading the power grid requires large-scale investments, yet money had been funneled into small, distributed sources like solar and wind, or dirty ones like coal and natural gas.

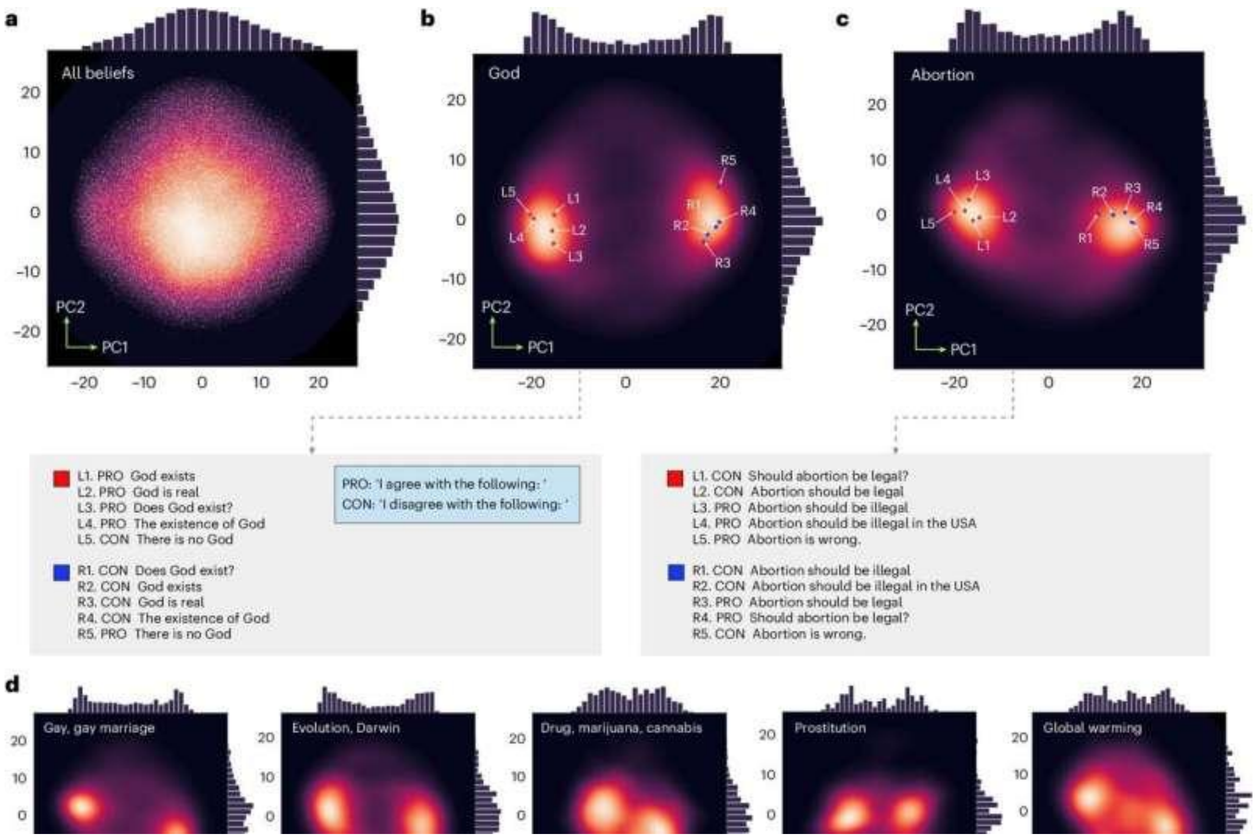
Now tech companies are making hundred-billion-dollar strategic bets that they can solve both nuclear dilemmas. They are betting that nuclear can offer the kind of steady, clean power their AI ambitions require.

This could be an unexpected positive consequence of AI: the revitalization of one of the safest and cleanest energy sources available to humankind.

JUNE 22, 2025

LLMs delve into online debates to create a detailed map of human beliefs

by Ingrid Fadelli, Phys.org edited by Lisa Lock, reviewed by Robert Egan



Structure of belief space via PCA. Credit: *Nature Human Behaviour* (2025). DOI: 10.1038/s41562-025-02228-z

Large language models (LLMs), such as the model underpinning the functioning of the well-known conversational platform ChatGPT, have proved to be very promising for summarizing and generating written texts. However, they could also be interesting tools for conducting research rooted in psychology, behavioral science and other scientific disciplines.

Researchers at Indiana University have recently used LLMs to study the intricate and nuanced landscape of human beliefs by analyzing debates between internet users on online platforms. Their proposed methodology, outlined in an article in *Nature Human Behaviour*, allowed them to create a detailed map of human beliefs, unveiling patterns hinting at polarization (i.e., extreme divisions between groups with opposing viewpoints) and cognitive dissonance (i.e., the discomfort felt when exposed to beliefs conflicting with our own).

"My fundamental research goal is to understand why people engage in certain behaviors, utilizing data and AI/NLP (Natural Language Processing)," Jisun An, senior author of the paper, told Phys.org.

"In pursuing this, I came to realize that beliefs lie at the heart of human actions, as they profoundly influence our decision-making and behaviors. Furthermore, I noticed that language embedding spaces effectively preserve semantic meaning, and that recent Large Language Models (LLMs) contain a vast amount of information about language, knowledge, and people."

Following the release of the first LLMs, An gradually became convinced that these advanced machine learning–based models could be used to study human beliefs and behavior. This was the primary inspiration behind her recent paper, which specifically analyzed beliefs expressed by people online.

"Our research proposes a novel methodology for constructing a 'belief embedding space' to understand the complex system of human beliefs," explained An. "Simply put, this method involves arranging countless individual beliefs on a continuous, high-dimensional map.

"In this space, each belief occupies a unique 'location,' with semantically similar or related beliefs positioned closer to each other, and contrasting beliefs placed further apart. Think of it like a map of ideas, where 'eating healthy is important' might be close to 'regular exercise improves well-being,' while 'junk food is fine every day' would be very far away."

An and her colleagues created their "human belief map" by fine-tuning S-BERT (Sentence-BERT), a specialized model for generating high-quality sentence embeddings and measuring semantic similarity. Despite being a relatively smaller model, S-BERT is widely adopted for practical applications due to its efficiency and effectiveness.

"While previous research often focused on specific topics or a limited number of beliefs, we leveraged the extensive language comprehension and vast knowledge embedded within LLMs to create a comprehensive map encompassing a much wider variety of beliefs," said An.

Visualization of the belief embedding space. Various beliefs are represented in a high-dimensional space and projected into two dimensions using the UMAP method. Credit: Lee at al
"This belief embedding space goes beyond mere classification; it provides a powerful foundation for quantitatively analyzing the complex interplay among beliefs and how individuals accept or reject new information (i.e., the decision-making process)."

Using their LLM-based methods, the researchers were able to analyze a vast and diverse range of beliefs, which were previously difficult to map collectively within a single space. In addition, they could numerically calculate the semantic similarity or distance between specific beliefs, unveiling complex relationships between beliefs that are difficult to uncover using traditional qualitative research methods.

"Even as new beliefs emerge or societal changes occur, our method allows for continuous updating and expansion of the belief map using LLMs, ensuring it reflects evolving societal dynamics," said An. "We showed that our methodology successfully utilizes LLMs to construct a sophisticated 'map' of human beliefs, or 'belief embedding.' This opens new avenues for systematically and quantitatively analyzing the intricate system of human beliefs and, in doing so, provides a foundation for qualitatively and systematically studying people's decision-making processes."

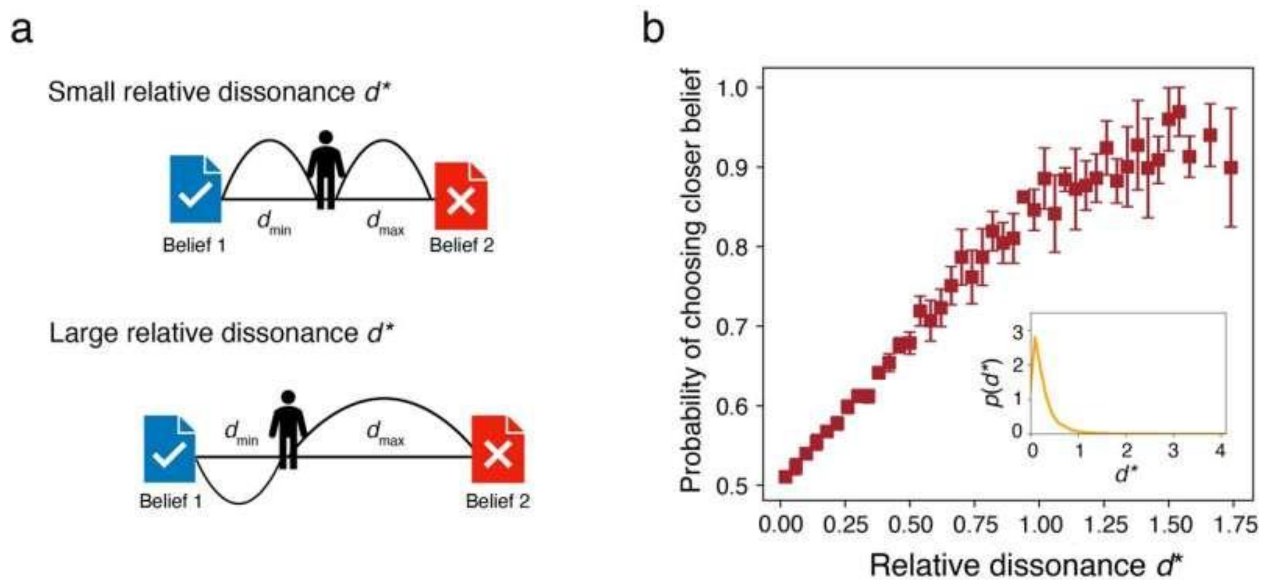
When they looked at the belief map created using LLMs, the researchers made some interesting discoveries. First, they found that "relative dissonance" significantly influences people's decision-making. This essentially means that when online users encounter new information or beliefs, they tend to choose or accept those that cause them less "discomfort" or are most aligned with their existing beliefs.

"More importantly, we show that people's belief choices are shaped not just by how close a belief is to their own, but by how much closer the belief is compared to its competing belief," explained An. "When two opposing beliefs on a certain issue are equally distant, people are just as likely to choose either one. However, when one belief is clearly closer than the other, people are far more likely to choose it."

An and her colleagues described the effect they observed when analyzing their belief map as "relative dissonance." This term essentially suggests that people's decisions are influenced by the relative gap between beliefs that are closer and further from their own. Specifically, the researchers found that the greater this gap is, the stronger a person's preference is for beliefs more aligned with their own.

"In other words, people are not only avoiding disagreement, but actively minimizing the difference in disagreement between available options," said An. "This finding highlights that decision-making is not driven by absolute distance alone, but by the relative discomfort of accepting a belief that feels much further away, echoing key ideas from cognitive dissonance theory."

The findings of this recent study could have various implications. First, they provide an explanation for why some information is readily accepted by some people and strongly rejected by others, shedding new light on the processes underpinning the formation and maintenance of social perspectives.



(a) Illustration of two scenarios in which a user selects a belief for a debate, highlighting contrasting cases of small and large relative dissonance, denoted as d*. When the two beliefs under consideration are equally distant from the user, selecting either belief results in a similar level of dissonance (small relative dissonance, small d*). In contrast, when one belief is significantly farther than the other, the potential dissonance experienced by the user depends on their selection (large relative dissonance, large d*). (b) The probability of choosing the closer belief increases linearly with d*. Credit: Lee at al

"Our work also offers guidance on how messages should be constructed to be effectively delivered to a target audience, by carefully considering their existing beliefs," said An. "It could also inform the refined design of policies or campaigns aimed at encouraging behavioral change in various fields, such as health or environmental initiatives, by better understanding the intricate interplay of beliefs."

The new insight gathered by An and her colleagues could contribute to the development of new behavioral science interventions aimed at encouraging people to make more responsible decisions that could benefit their health, finances or the environment on Earth. Meanwhile, the researchers plan to continue using LLMs to study people's beliefs and online behavior.

"While our current study utilized limited data from Debate.org (DDO) to construct the belief map, our immediate project involves leveraging larger and more diverse social media datasets, such as Reddit, to build an even more detailed, richer, and real-world reflective belief map," said An. "This will enable us to capture more subtle differences in individual beliefs and analyze belief interactions in various contexts more accurately."

Once they have created this more refined map of human beliefs, the researchers plan to use it to plan new studies and experiments. They would also like to connect their observations with the results of another project carried out at Indiana University, called BRAIN (Belief Resonance and AI Narratives).

"This new project will delve deeper into how an individual's belief system interacts with new incoming information and the mechanisms by which that information is either accepted or rejected," said An. "For example, we want to understand why a community might quickly embrace a new sustainable farming practice, while another, with different core beliefs about traditional methods, might strongly resist it."

So far, An and her colleagues have used LLMs to analyze people's comments and posts on popular social media platforms. In the future, they would also like to explore the relationship between the beliefs that people express online and their decision to join or leave specific online communities.

"We hope that these further studies will allow us to better understand how beliefs are connected to people's actual behaviors, including social interactions and decision-making," added An. "We believe this will make significant contributions to understanding and predicting various social phenomena."

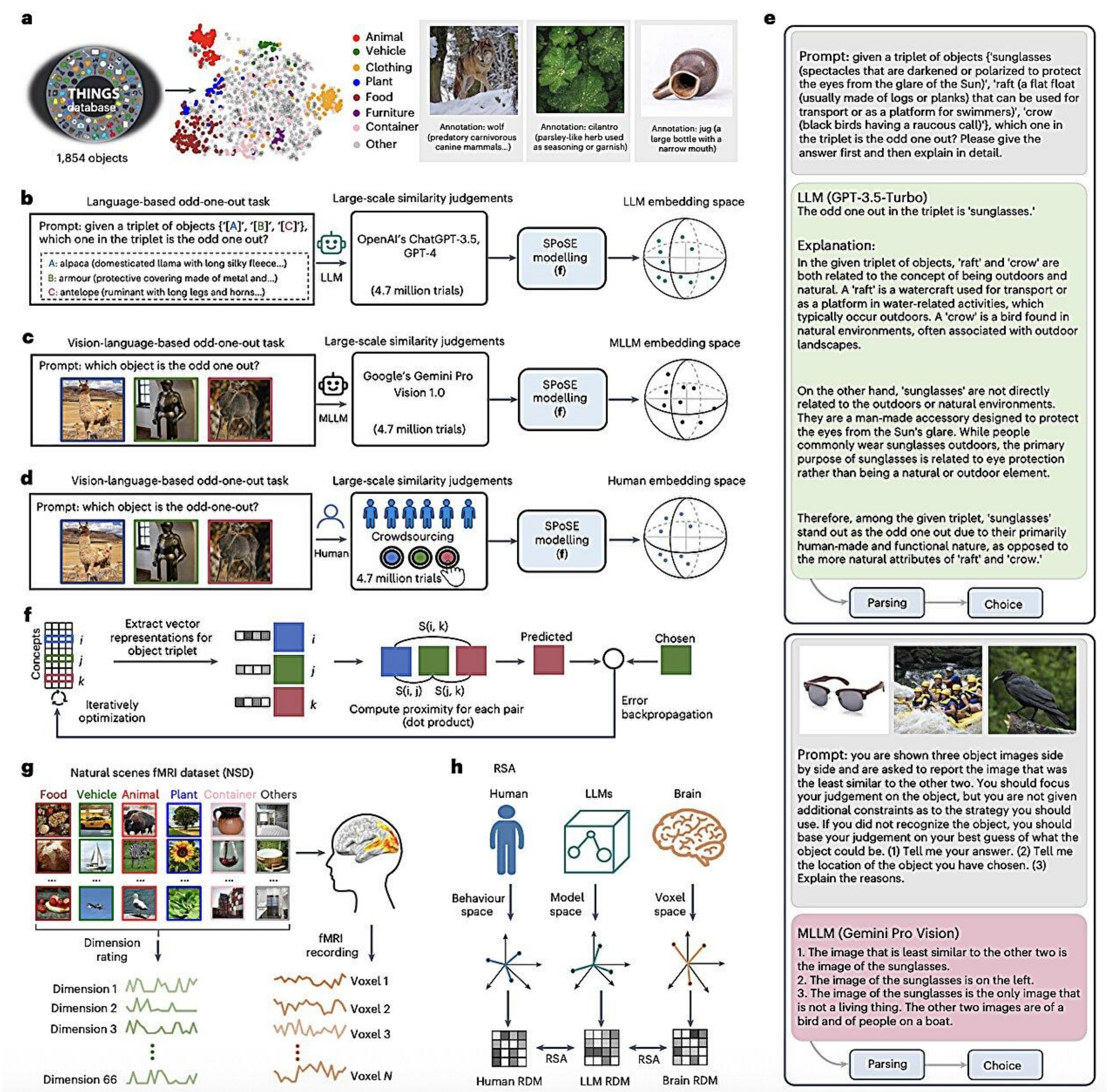
Written for you by our author Ingrid Fadelli, edited by Lisa Lock , and fact-checked and reviewed by Robert Egan —this article is the result of careful human work. We rely on readers like you to keep independent science journalism alive. If this reporting matters to you, please consider a donation (especially monthly). You'll get an **ad-free** account as a thank-you.

More information: Byunghwee Lee et al, A semantic embedding space based on large language models for modelling human beliefs, *Nature Human Behaviour* (2025). DOI: 10.1038/s41562-025-02228-z
Journal information: Nature Human Behaviour

JUNE 25, 2025

Multimodal LLMs and the human brain create object representations in similar ways, study finds

by Ingrid Fadelli, Phys.org edited by Lisa Lock, reviewed by Robert Egan



Schematics of the experiment and analysis methods. a, THINGS database and examples of object images with their language descriptions given at the bottom. b–d, Pipelines of mental embedding learning under the triplet odd-one-out paradigm for LLMs (b), MLLMs (c) and humans (d). e, Examples of prompts and responses for LLMs and MLLMs. f, Illustration of the SPoSE modelling approach. g, Illustration of the NSD dataset with dimension ratings for stimulus images. h, Overview of the comparisons between space of LLMs, human behaviour and brain activity. All images were replaced with similar images from Pixabay and Pexels under a Creative Commons license CC0. Credit: *Nature Machine Intelligence* (2025). DOI: 10.1038/s42256-025-01049-z

A better understanding of how the human brain represents objects that exist in nature, such as rocks, plants, animals, and so on, could have interesting implications for research in various fields, including psychology, neuroscience and computer science. Specifically, it could help shed new light on how humans interpret sensory information and complete different real-world tasks, which could also inform the development of artificial intelligence (AI) techniques that closely emulate biological and mental processes.

Multimodal large language models (LLMs), such as the latest models underpinning the functioning of the popular conversational platform ChatGPT, have been found to be highly effective computational techniques for the analysis and generation of texts in various human languages, images and even short videos.

As the texts and images generated by these models are often very convincing, to the point that they could appear to be human-created content, multimodal LLMs could be interesting experimental tools for studying the underpinnings of object representations.

Researchers at the Chinese Academy of Sciences recently carried out a study aimed at better understanding how multimodal LLMs represent objects, while also trying to determine whether the object representations that emerge in these models resemble those observed in humans. Their findings are published in *Nature Machine Intelligence*.

"Understanding how humans conceptualize and categorize natural objects offers critical insights into perception and cognition," Changde Du, Kaicheng Fu and their colleagues wrote in their paper. "With the advent of large language models (LLMs), a key question arises: Can these models develop human-like object representations from linguistic and multimodal data?

"We combined behavioral and neuroimaging analyses to explore the relationship between object concept representations in LLMs and human cognition."

Object dimensions illustrating their interpretability. Credit: *Nature Machine Intelligence* (2025). DOI: 10.1038/s42256-025-01049-z

As part of their study, the researchers specifically examined the object representations emerging in the LLM ChatGPT- 3.5 created by Open AI, and in the multi-modal LLM GeminiPro Vision 1.0 developed at Google DeepMind. They asked these models to complete simple tasks known as triplet judgments. For each of these tasks, they were presented with three objects and asked to select the two that more closely resembled each other.

"We collected 4.7 million triplet judgments from LLMs and multimodal LLMs to derive low-dimensional embeddings that capture the similarity structure of 1,854 natural objects," wrote Du, Fu and their colleagues. "The resulting 66-dimensional embeddings were stable, predictive and exhibited semantic clustering similar to human mental representations. Remarkably, the dimensions underlying these embeddings were interpretable, suggesting that LLMs and multimodal LLMs develop human-like conceptual representations of objects."

Using the large dataset of triplet judgments that they collected, the researchers computed low-dimensional embeddings. These are mathematical representations that outline the similarity between objects over various dimensions, placing similar objects closer to each other in an abstract space.

Notably, the researchers observed that the low-dimensional embeddings they attained reliably grouped objects into meaningful categories, such as "animals," "plants," and so on. They thus concluded that LLMs and multi-modal LLMs naturally organize objects similarly to how they are represented and categorized in the human mind.

"Further analysis showed strong alignment between model embeddings and neural activity patterns in brain regions such as the extra-striate body area, para-hippocampal place area, retro-splenial cortex and fusiform face area," the team wrote. "This provides compelling evidence that the object representations in LLMs, although not identical to human ones, share fundamental similarities that reflect key aspects of human conceptual knowledge."

Overall, the results gathered by Du, Fu and their colleagues suggest that human-like natural object representations could inherently emerge in LLMs and multi-modal LLMs after they are trained on large amounts of data. In the future, this study could inspire other research teams to explore how LLMs represent objects, while also potentially contributing to the further advancement of brain-inspired AI systems.

Written for you by our author Ingrid Fadelli, edited by Lisa Lock , and fact-checked and reviewed by Robert Egan —this article is the result of careful human work. We rely on readers like you to keep independent science journalism alive. If this reporting matters to you, please consider a donation (especially monthly). You'll get an **ad-free** account as a thank-you.

More information: Changde Du et al, Human-like object concept representations emerge naturally in multimodal large language models, *Nature Machine Intelligence* (2025). DOI: 10.1038/s42256-025-01049-z
Journal information: Nature Machine Intelligence

DeepMind’s new AlphaGenome AI tackles the ‘dark matter’ in our DNA 25 June 2025

Tool aims to solve the mystery of non-coding sequences — but is still in its infancy.

Researchers feed vast quantities of genomic data into machine-learning systems to train them to predict the role of non-coding sequences.Credit: JuSun/iStock via Getty

Nearly 25 years after scientists completed a draft human genome sequence, many of its 3.1 billion letters remain a puzzle. The 98% of the genome that is not made of protein-coding genes — but which can influence their activity — is especially vexing.

An artificial intelligence (AI) model developed by Google DeepMind in London could help scientists to make sense of this ‘dark matter’, and see how it might contribute to diseases such as cancer and influence the inner workings of cells. The model, called AlphaGenome, is described in a 25 June preprint.

“This is one of the most fundamental problems not just in biology — in all of science,” said Pushmeet Kohli, the company’s head of AI for science, at a press briefing.

The ‘sequence to function’ model takes long stretches of DNA and predicts various properties, such as the expression levels of the genes they contain and how those levels could be affected by mutations.

“I think it is an exciting leap forward,” says Anshul Kundaje, a computational genomicist at Stanford University in Palo Alto, California, who has had early access to AlphaGenome. “It is a genuine improvement in pretty much all current state-of-the-art sequence-to-function models.”

An ‘all in one’ approach

When DeepMind unveiled AlphaFold 2 in 2020, it went a long way towards solving a problem that had challenged researchers for decades: determining how a protein’s sequence contributes to its 3D shape.

Working out what DNA sequences do is different, because there is no one answer, as in a 3D structure that AlphaFold delivers. A single DNA stretch will have numerous, interconnected roles — from attracting one set of cellular machinery to latch onto a particular section of a chromosome and turn a nearby gene into an RNA molecule, to attracting protein-transcription factors that influence where, when and to what extent gene expression occurs. Many DNA sequences, for example, influence gene activity by altering a chromosome’s 3D shape, either restricting or easing access for the machinery that does the transcription.

Biologists have been chipping away at this question for decades with various kinds of computational tool. In the past decade or so, scientists have developed dozens of AI models to make sense of the genome. Many of these have focused on an individual task, such as predicting levels of gene expression or determining how modular segments of individual genes, called exons, are cut-and-pasted into distinct proteins. But scientists are increasingly interested in ‘all in one’ tools for interpreting DNA sequences.

AlphaGenome is one such model. It can take inputs of up to one million DNA letters — a stretch that could include a gene and myriad regulatory elements — and make thousands of predictions about numerous biological properties. In many cases, AlphaGenome’s predictions are sensitive to single-DNA-letter changes, which means that scientists can predict the consequences of mutations.

In one example, DeepMind researchers applied the AlphaGenome model to diverse mutations identified in previous studies in people with a type of leukaemia. The model accurately predicted that the non-coding mutations indirectly activated a nearby gene that is a common driver of this cancer.

WHAT THE HUMAN BRAIN SEES THAT AI CAN'T: NEW STUDY REVEALS OUR UNIQUE EDGE IN NAVIGATING THE WORLD

TIM MCMILLAN·JUNE 23, 2025

Most of us have probably never paused to consider the awe-inspiring complexity of our brain's ability to effortlessly decide whether to walk, climb, or swim when encountering a new environment. In an instant, without conscious thought, our brains size up a scene and tell us how we can move through it, a feat that even the most advanced AI systems struggle to replicate.

However, according to new research, this remarkable ability we take for granted daily relies on specialized neural processes that artificial intelligence has yet to replicate despite its rapid advances.

Researchers from the University of Amsterdam have uncovered how the human brain encodes so-called "locomotive action affordances"—the opportunities for movement that our surroundings present.

Published in *PNAS Neuroscience*, their work provides fresh insights into how our brains instantly recognize whether we can walk through a field, climb over rocks, or dive into water, all at a glance.

Perhaps most strikingly, the study highlights how deep neural networks (DNNs)—AI systems inspired by biological brains—fail to replicate this fundamental aspect of human perception.

The team's findings point to a profound difference between natural and artificial intelligence, with implications for everything from robotics to self-driving cars. While AI systems have made enormous strides in recognizing objects and classifying scenes, they still fail to understand what those scenes afford regarding physical action.

"Even the best AI models don't give exactly the same answers as humans, even though it's such a simple task for us," study co-author and computational neuroscientist Dr. Iris Groen explained in a press release. “This shows that our way of seeing is deeply intertwined with how we interact with the world.”

“We connect our perception to our experience in a physical world. AI models can’t do that because they only exist in a computer.”

The researchers combined brain imaging, behavioral studies, and machine learning analysis to explore this capability. Volunteers in the study were shown images of indoor and outdoor environments, and their brain activity was monitored using functional MRI (fMRI).

Sample of images shown during experiments. (Image Source: University of Amsterdam, Bartnik, et al.)

The goal was to map how the brain represents different locomotive possibilities, such as walking, climbing, swimming, crawling, jumping, or flying.

“We wanted to know: when you look at a scene, do you mainly see what is there – such as objects or colors – or do you also automatically see what you can do with it,” Dr. Groen explained. “Psychologists call the latter ‘affordances’ – opportunities for action; imagine a staircase that you can climb or an open field that you can run through.”

The data revealed that specific regions of the human visual cortex—especially those areas responsible for processing scenes—light up in patterns that directly encode the types of movement possible in a given environment.

Intriguingly, these patterns were distinct from those activated by other properties, such as recognizing objects, surfaces, or general scene categories. This suggests that our brain carves out a separate space to understand how we can act within our surroundings.

In parallel, the researchers tested a variety of deep neural networks trained on everyday visual tasks like object recognition or scene classification. These AIs could identify whether an image depicted a kitchen, a forest, or a city street—but they didn’t do well at perceiving what movements those environments allowed.

Results showed that the alignment between AI activations and human brain patterns was weak in locomotive affordances, even though it was strong for object-related tasks.

The team then tried fine-tuning these AI models, training them to classify images by affordance or using language embeddings focused on action possibilities.

While this improved the alignment somewhat, none of the tested systems fully captured the nuanced way human brains represent locomotive possibilities. The gap between human and machine perception remains wide, at least in this domain.

“When trained specifically for action recognition, they could somewhat approximate human judgments, but the human brain patterns didn’t match the models’ internal calculations,” Dr. Groen said.

One of the study’s key takeaways is that understanding how we navigate the world requires more than recognizing what’s in it—it requires grasping what we can do within it. Our brains manage this automatically, without conscious thought and without needing explicit instructions.

“Our results suggest that locomotive action affordance perception in scenes relies on specialized neural representations different from those used for other visual understanding tasks,” the researchers wrote. “Training DNNs directly on affordance labels or using affordance-centered language embeddings increases alignment with human behavior, but none of the tested models fully captures locomotive action affordance perception.”

The study opens new and exciting avenues for improving artificial intelligence systems, particularly those designed to operate in dynamic environments. Better models of affordance perception could revolutionize self-driving vehicles, delivery robots, and even AI assistants in virtual environments, allowing them to interact with the world in more human-like ways. The potential impact of this research on AI’s future is hopeful and exciting.

For now, the work underscores how much we still have to learn from our own biology. AI may have surpassed human capabilities in specific tasks like playing chess or analyzing massive data sets.

However, we remain far ahead when it comes to the intuitive understanding of space and movement—something that allows a child to clamber over playground equipment or an adult to navigate a crowded city street. This underscores the need for further research and development in AI to bridge this gap.

The results also remind us that intelligence is more than data processing or pattern matching. The human brain sees possibilities for action, weaving together perception and potential in ways that today's artificial minds are only beginning to grasp.

Studies like this will be essential as AI researchers work to bridge this gap. Scientists hope these new insights will help inspire the next generation of AI systems—ones that can navigate the world with greater efficiency and energy savings, much like the human brain.

"Current AI training methods use a huge amount of energy and are often only accessible to large tech companies," Dr. Groen said. "More knowledge about how our brain works, and how the human brain processes certain information very quickly and efficiently, can help make AI smarter, more economical, and more human-friendly."

MOST COUNTRIES ARE FALLING BEHIND IN PREPARING THE WORKFORCE FOR AI, NEW STUDY WARNS

TIM MCMILLAN·JUNE 23, 2025

As artificial intelligence (AI) rapidly reshapes the modern workplace, most of the world is alarmingly unprepared to help its workers keep pace.

A new study by researchers at the University of Georgia finds that despite the growing dominance of AI technologies, only a small fraction of countries have meaningful strategies to ensure their workforce is ready for this seismic shift.

The study, published in *Human Resource Development Review*, delves into the importance of comprehensive national AI strategies (NAISs) in preparing the workforce for the AI revolution. It analyzed the NAISs of 50 countries, specifically focusing on their education and workforce development policies.

The researchers found that just 13 nations—mostly developed economies in Europe—demonstrate high-level prioritization of AI readiness through clear objectives, detailed plans, and robust support measures. The rest, the study suggests, are either treading water or, worse, failing to act at all.

"Advancements in artificial intelligence signal a significant shift toward more automated and data-driven workplaces," the study authors write. "Emphasizing the need for Human Resource Development (HRD) to prepare the workforce with adequate AI competencies for AI-empowered environments."

From logistics and manufacturing to healthcare, finance, and even education, AI tools are transforming how jobs are performed and redefining the skills employees need to thrive. As technology barrels forward, the need for public policy and workforce planning to catch up is becoming more apparent.

The University of Georgia research team undertook a comprehensive review of how national governments are responding to this challenge through formal AI strategies. They aimed to pinpoint which policies and initiatives are being used to guide human resource development (HRD) efforts in this new landscape.

The results revealed that only 13 countries have made AI workforce development a national priority. Eleven of those countries were in Europe, with Australia and Mexico being the two exceptions.

Researchers say these countries have set clear goals backed up with specific measures to implement educational and training strategies. They are investing in programs that reskill existing workers, cultivate AI talent pipelines, and build support systems that help ensure success.

By contrast, most countries surveyed have yet to outline a coherent plan for preparing their workforce. Their national AI strategies either fail to mention workforce training in detail or provide only vague statements of intent without clear pathways for action.

In other words, while AI is poised to upend global labor markets, most nations remain at the starting line in preparing their citizens for what's ahead.

The study also identified six key categories of educational and training strategies designed to bolster AI talent preparation and workforce reskilling. Although these strategies vary by country, they generally focus on integrating AI-related content into academic curricula, supporting lifelong learning initiatives, encouraging public-private partnerships in training efforts, and developing specialized programs for technical and non-technical workers.

The researchers also highlighted four types of support resources—such as financial investments, technology infrastructure, partnerships, and policy frameworks—that are essential to making these strategies effective.

One of the study's central messages is that preparing the workforce for AI is not just about technical training. It's about designing HRD curricula, programs, and policies that are workplace-oriented, inclusive, and sensitive to each country's unique contextual and cultural factors.

What works in a European nation with a highly developed education system and robust social safety nets may not translate directly to a country with different resources, priorities, or challenges.

Importantly, the study calls for more research on effectively training workers for AI-powered environments. The authors emphasize the need for new theories and models of AI workforce education and innovative pedagogical strategies that reflect the realities of today's and tomorrow's workplaces.

Without clear strategies, countries risk leaving large segments of their workforce behind as AI tools become increasingly central to everyday work. This could exacerbate inequality, stifle economic growth, and create friction in societies unprepared for the disruption that AI will inevitably bring.

It's not enough for nations to embrace AI in business and industry. They must also ensure that their citizens are equipped to work alongside these technologies—ideally, to help shape their future applications.

For example, national conversations about AI often focus on competitiveness, ethics, or security concerns in the United States. However, when it comes to preparing the workforce, the U.S. lags behind some of its European peers in translating AI priorities into detailed plans for education and training. This leaves businesses and local governments to fill the gaps, leading to fragmented efforts that may not reach all communities equally.

The United States was among 23 countries that treated workforce training for AI as a medium priority, offering few detailed plans for AI education. Perhaps surprisingly, given the intense technological competition between Washington and Beijing, China was rated as placing a low priority on AI education.

Meanwhile, countries like Germany and Italy are setting the pace with national initiatives emphasizing digital literacy from an early age.

However, while some governments have taken various steps to promote technical training and digital literacy, the study found that few countries explicitly prioritize developing uniquely human “soft skills,“ potentially leaving a significant gap in efforts to future-proof the workforce.

“Human soft skills, such as creativity, collaboration, and communication, cannot be replaced by AI,“ study co-author and an assistant research scientist at the UGA, Dr. Lehong Shi, said in a press release. “And they were only mentioned by a few countries.“

It’s clear that AI’s rise is inevitable. However, the question is whether nations will rise to the challenge of preparing their people for this new reality—or watch AI technology reshape the workforce without them.

As the researchers conclude, building an AI-competent workforce is one of the defining challenges of our time. For policymakers, educators, and business leaders, the clock is ticking.

“AI skills and competencies are very important,“ Dr. Shi explained. “If you want to be competitive in other areas, it’s very important to prepare employees to work with AI in the future.”

How Generative Engine Optimization (GEO) Rewrites the Rules of Search

Zach Cohen and **Seema Amble** May 28, 2025

It’s the end of search as we know it, and marketers feel fine. Sort of.

For over two decades, SEO was the default playbook for visibility online. It spawned an entire industry of keyword stuffers, backlink brokers, content optimizers, and auditing tools, along with the professionals and agencies to operate them. But in 2025, search has been shifting away from traditional browsers toward LLM platforms. With Apple’s announcement that AI-native search engines like Perplexity and Claude will be built into Safari, Google’s distribution chokehold is in question. The foundation of the $80 billion+ SEO market just cracked.

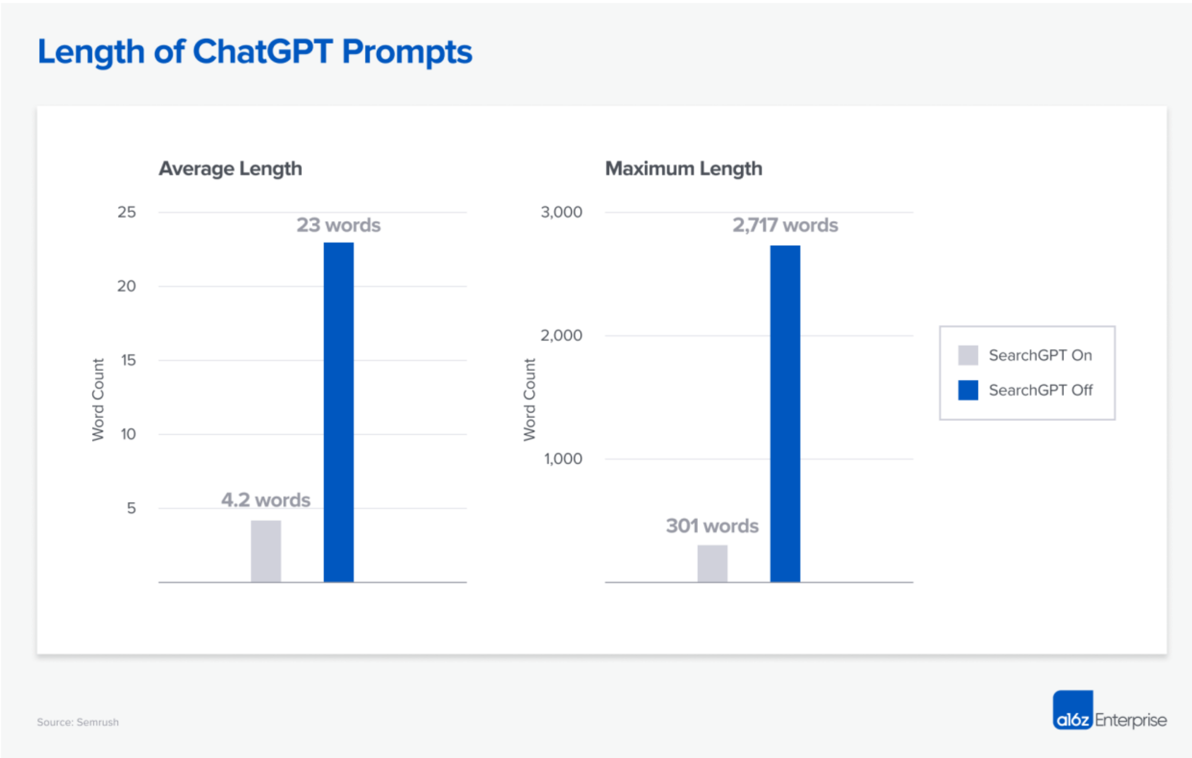
A new paradigm is emerging, one driven not by page rank, but by language models. We’re entering Act II of search: Generative Engine Optimization (GEO).

From links to language models

Traditional search was built on links. GEO is built on language.

In the SEO era, visibility meant ranking high on a results page. Page ranks were determined by indexing sites based on keyword matching, content depth and breadth, backlinks, user experience engagement, and more. Today, with LLMs like GPT-4o, Gemini, and Claude acting as the interface for how people find information, visibility means showing up directly in the answer itself, rather than ranking high on the results page.

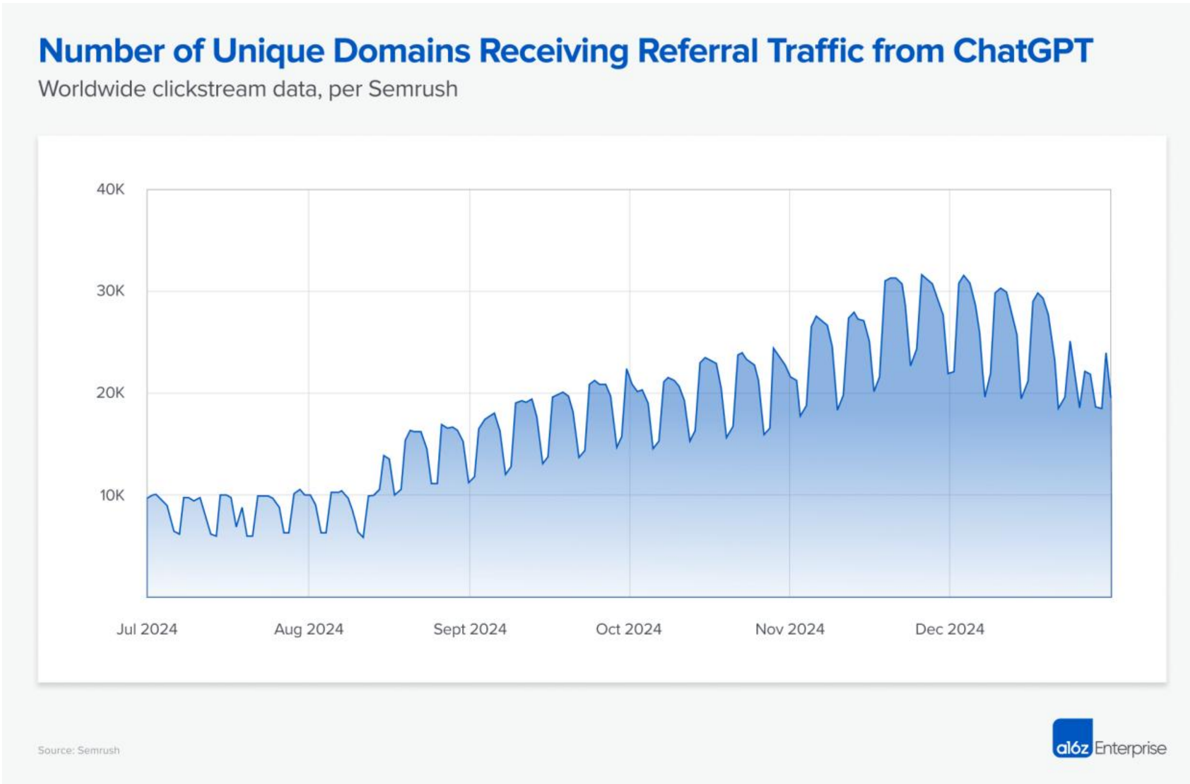
As the format of the answers changes, so does the way we search. AI-native search is becoming fragmented across platforms like Instagram, Amazon, and Siri, each powered by different models and user intents. Queries are longer (23 words, on average, vs. 4), sessions are deeper (averaging 6 minutes), and responses vary by context and source. Unlike traditional search, LLMs remember, reason, and respond with personalized, multi-source synthesis. This fundamentally changes how content is discovered and how it needs to be optimized.



Traditional SEO rewards precision and repetition; generative engines prioritize content that is well-organized, easy to parse, and dense with meaning (not just keywords). Phrases like “in summary” or bullet-point formatting help LLMs extract and reproduce content effectively.

It’s also worth noting that the LLM market is also fundamentally different from the traditional search market in terms of business model and incentives. Classic search engines like Google monetized user traffic through ads; users paid with their data and attention. In contrast, most LLMs are paywalled, subscription-driven services. This structural shift affects how content is referenced: there’s less of an incentive by model providers to surface third-party content, unless it’s additive to the user experience or reinforces product value. While it’s possible that an ad market may eventually emerge on top of LLM interfaces, the rules, incentives, and participants would likely look very different than traditional search.

In the meantime, one emerging signal of the value in LLM interfaces is the volume of outbound clicks. ChatGPT, for instance, is already driving referral traffic to tens of thousands of distinct domains.



From rankings to model relevance

It's no longer just about click-through rates, it's about reference rates: how often your brand or content is cited or used as a source in model-generated answers. In a world of AI-generated outputs, GEO means optimizing for what the model *chooses* to reference, not just whether or where you appear in traditional search. That shift is revamping how we define and measure brand visibility and performance.

Already, new platforms like **Profound**, **Goodie**, and **Daydream** enable brands to analyze how they appear in AI-generated responses, track sentiment across model outputs, and understand which publishers are shaping model behavior. These platforms work by fine-tuning models to mirror brand-relevant prompt language, strategically injecting top SEO keywords, and running synthetic queries at scale. The outputs are then organized into actionable dashboards that help marketing teams monitor visibility, messaging consistency, and competitive share of voice.

Canada Goose used one such tool to gain insight into how LLMs referenced the brand — not just in terms of product features like warmth or waterproofing, but brand recognition itself. The takeaways were less about how users discovered Canada Goose, but whether the model spontaneously mentioned the brand at all, an indicator of unaided awareness in the AI era.

This kind of monitoring is becoming as important as traditional SEO dashboards. Tools like **Ahrefs**' Brand Radar now track brand mentions in AI Overviews, helping companies understand how they're framed and remembered by generative engines. **Semrush** also has a dedicated AI toolkit designed to help brands track perception across generative platforms, optimize content for AI visibility, and respond quickly to emerging mentions in LLM outputs, a sign that legacy SEO players are adapting to the GEO era.

We're seeing the emergence of a new kind of brand strategy: one that accounts not just for perception in the public, but perception in the model. How you're encoded into the AI layer is the new competitive advantage.

Of course, GEO is still in its experimental phase, much like the early days of SEO. With every major model update, we risk relearning (or unlearning) how to best interact with these systems. Just as Google's search algorithm updates once caused companies to scramble to counter fluctuating rankings, LLM providers are still tuning the rules behind what their models cite. Multiple schools of thought are emerging: some GEO tactics are fairly well understood (e.g., being mentioned in source documents LLMs cite), while other assumptions are more speculative, such as whether models prioritize journalistic content over social media, or how preferences shift with different training sets.

Lessons from the SEO era

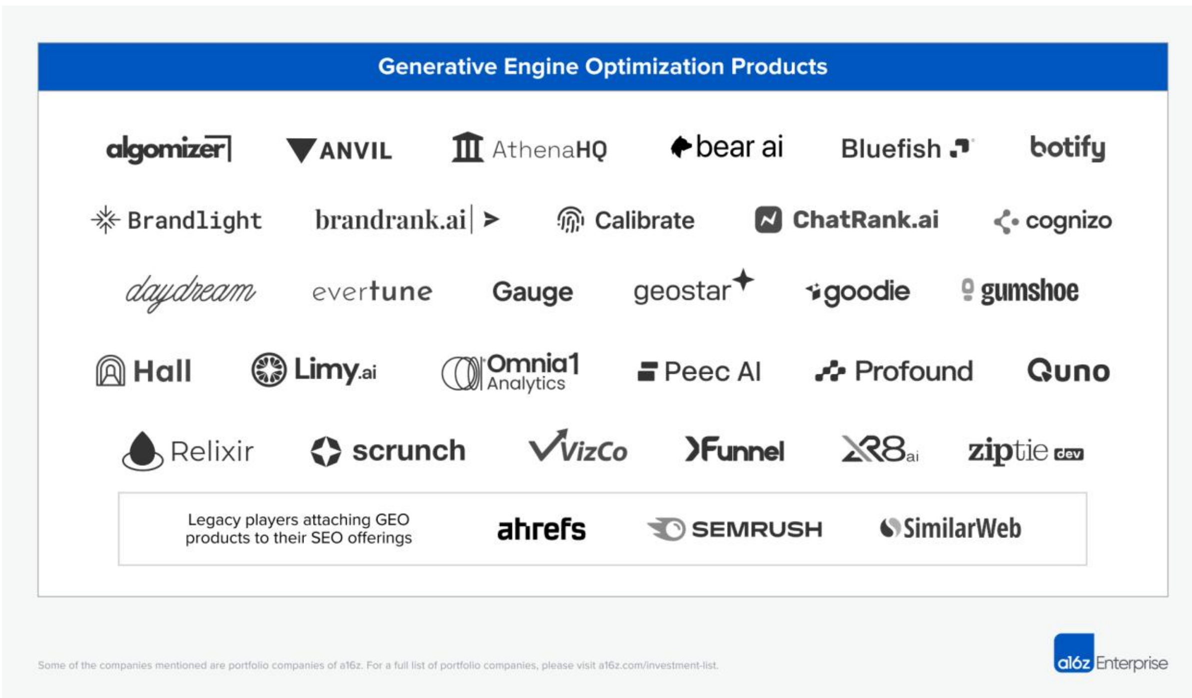
Despite its scale, SEO never produced a monopolistic winner. Tools that helped companies with SEO and keyword research, like **Semrush**, **Ahrefs**, **Moz**, and **Similarweb**, were successful in their own right, but none captured the full stack (or grew via acquisition, like Similarweb). Each carved out a niche: backlink analysis, traffic monitoring, keyword intelligence, or technical audits.

SEO was always fragmented. The work was distributed across agencies, internal teams, and freelance operators. The data was messy and rankings were inferred, not verified. Google held the algorithmic keys, but no vendor ever controlled the interface. Even at its peak, the biggest SEO players were tooling providers. They didn't have the user engagement, data control, or network effects to become hubs where SEO activity is concentrated. Clickstream data — records of the links users click as they navigate websites — is arguably the clearest window into real user behavior. Historically, though, this data has been prohibitively hard to access, locked behind

ISPs, SDKS, browser extensions, and data brokers. This made building accurate, scalable insights nearly impossible without deep infrastructure or privileged access.

GEO changes that.

How to make the mentions: The emergence of GEO tools



This isn't just a tooling shift, it's a platform opportunity. The most compelling GEO companies won't stop at measurement. They'll fine-tune their own models, learning from billions of implicit prompts across verticals. They'll own the loop — insight, creative input, feedback, iteration — with differentiated technology that doesn't just observe LLM behavior, but shapes it. They'll also figure out a way to capture clickstream data and combine first- and third-party data sources.

Platforms that win in GEO will go beyond brand analysis and provide the infrastructure to act: generating campaigns in real time, optimizing for model memory, and iterating daily, as LLM behavior shifts. These systems will be operational.

That unlocks a much broader opportunity than visibility. If GEO is how a brand ensures it's referenced in AI responses, it's also how it manages its ongoing relationship with the AI layer itself. GEO becomes the system of record for interacting with LLMs, allowing brands to track presence, performance, and outcomes across generative platforms. Own that layer, and you own the budget behind it.

That's the monopolistic potential: not just serving insights, but becoming *the* channel. If SEO was a decentralized, data-adjacent market, GEO can be the inverse — centralized, API-driven, and embedded directly into brand workflows. Ultimately, GEO by itself is perhaps the most obvious wedge, especially as we see a shift in search behavior, but ultimately, it's really a wedge into performance marketing, more broadly. The same brand guidelines and understanding of user data that power GEO can power growth marketing. This is how a big business gets built, as a software product is able to test multiple channels, iterate, and optimize across them. AI enables an autonomous marketer.

Timing matters. Search is just beginning to shift, but ad dollars move fast, especially when there's arbitrage. In the 2000s, that was Google's Adwords. In the 2010s, it was Facebook's targeting engine. Now, in 2025, it's LLMs and the platforms that help brands navigate how their content is ingested and referenced by those models. Put another way, GEO is the competition to get into the model's mind.

In a world where AI is the front door to commerce and discovery, the question for marketers is: Will the model remember you?

The Dream of an AI Scientist Is Closer Than Ever

The number of scientific papers relying on AI has quadrupled, and the scope of problems AI can tackle expands by the day. Jun 26, 2025

Modern artificial intelligence is a product of decades of painstaking scientific research. Now, it's starting to pay that effort back by accelerating progress across academia.

Ever since the emergence of AI as a field of study, researchers have dreamed of creating tools smart enough to accelerate humanity's endless drive to acquire new knowledge. With the advent of deep learning in the 2010s, this goal finally became a realistic possibility.

Between 2012 and 2022, the proportion of scientific papers that have relied on AI in some way has quadrupled to almost 9 percent. Researchers are using neural networks to analyze data, conduct literature reviews, or model complex processes across every scientific discipline. And as the technology advances, the scope of problems they can tackle is expanding by the day.

The poster boy for AI's use in science is undoubtedly Google DeepMind's Alphafold, whose inventors won the 2024 Nobel Prize in Chemistry. The model used advances in transformers—the architecture that powers large language models—to solve the "protein folding problem" that had bedeviled scientists for decades.

A protein's structure determines its function, but previously the only way to discover its shape was with complex imaging techniques like X-ray crystallography and cryo-electron microscopy. Alphafold, in comparison, could predict the shape of a protein from nothing more than the series of amino acids making it up, something computer scientists had been trying and failing to do for years.

This made it possible to predict the shape of every protein known to science in just two years, a feat that could have transformative impact on biomedical research. Alphafold 3, released in 2024, goes even further. It can predict both the structure and interactions of proteins, as well as DNA, RNA, and other biomolecules.

Google has also turned its AI loose on another area of the life sciences, working with Harvard researchers to create the most detailed map of human brain connections to date. The team took ultra-thin slices from a 1-millimeter cube of human brain and used AI-based imaging technology to map the roughly 50,000 cells and 150 million synaptic connections within.

This is by far the most detailed "connectome" of the human brain produced to date, and the data is now freely available, providing scientists a vital tool for exploring neuronal architecture and connectivity. This could boost our understanding of neurological disorders and potentially provide insights into core cognitive processes like learning and memory.

AI is also revolutionizing the field of materials science. In 2023, Google DeepMind released a graph neural network called GnoME that predicted 2.2 million novel inorganic crystal structures, including 380,000 stable ones that could potentially form the basis of new technologies.

Not to be outdone, other big AI developers have also jumped into this space. Last year, Meta released and open sourced its own transformer-based materials discovery models and, crucially, a dataset with more than 110 million materials simulations that it used to train them, which should allow other researchers to build their own materials science AI models.

Earlier this year Microsoft released MatterGen, which uses a diffusion model—the same architectures used in many image and video generation models—to produce novel inorganic crystals. After fine-tuning, they showed it could be prompted to produce materials with specific chemical, mechanical, electronic, and magnetic properties.

One of AI's biggest strengths is its ability to model systems far too complex for conventional computational techniques. This makes it a natural fit for weather forecasting and climate modeling, which currently rely on enormous physical simulations running on supercomputers. Google DeepMind's GraphCast model was the first to show the promise of the approach, which used graph neural networks to generate 10-day forecasts in one minute and at higher accuracy than existing gold standard approaches that would take several hours.

AI’s Ability to See the Future

The brain quickly adapts to change by predicting multiple futures, neuron by neuron. These finding could lead to AI that can do the same thing. Jun 06, 2025

We constantly make decisions. Some seem simple: I booked dinner at a new restaurant, but I'm hungry now. Should I grab a snack and risk losing my appetite or wait until later for a satisfying meal—in other words, what choice is likely more rewarding?

Dopamine neurons inside the brain track these decisions and their outcomes. If you regret a choice, you'll likely make a different one next time. This is called reinforcement learning, and it helps the brain continuously adjust to change. It also powers a family of AI algorithms that learn from successes and mistakes like humans do.

But reward isn't all or nothing. Did my choice make me ecstatic, or just a little happier? Was the wait worth it?

This week, researchers at the Champalimaud Foundation, Harvard University, and other institutions said they've discovered a previously hidden universe of dopamine signaling in the brain. After recording the activity of single dopamine neurons as mice learned a new task, the teams found the cells don't simply track rewards. They also keep tabs on when a reward came and how big it was—essentially building a mental map of near-term and far-future reward possibilities.

"Previous studies usually just averaged the activity across neurons and looked at that average," said study author Margarida Sousa in a press release. "But we wanted to capture the full diversity across the population—to see how individual neurons might specialize and contribute to a broader, collective representation."

Some dopamine neurons preferred immediate rewards; others slowly ramped up activity in expectation of delayed satisfaction. Each cell also had a preference for the size of a reward and listened out for internal signals—for example, if a mouse was thirsty, hungry, and its motivation level.

Surprisingly, this multidimensional map closely mimics some emerging AI systems that rely on reinforcement learning. Rather than averaging different opinions into a single decision, some AI systems use a group of algorithms that encodes a wide range of reward possibilities and then votes on a final decision.

In several simulations, AI equipped with a multidimensional map better handled uncertainty and risk in a foraging task.

The results "open new avenues" to design more efficient reinforcement learning AI that better predicts and adapts to uncertainties, wrote one team. They also provide a new way to understand how our brains make everyday decisions and may offer insight into how to treat impulsivity in neurological disorders such as Parkinson's disease.

Dopamine Spark

For decades, neuroscientists have known dopamine neurons underpin reinforcement learning. These neurons puff out a small amount of dopamine—often dubbed the pleasure chemical—to signal an unexpected reward. Through trial and error, these signals might eventually steer a thirsty mouse through a maze to find the water stashed at its end. Scientists have developed a framework for reinforcement learning by recording the electrical activity of dopamine neurons as these critters learned. Dopamine neurons spark with activity in response to nearby rewards, then this activity slowly fades as time goes by—a process researchers call "discounting."

But these analyses average activity into a single expected reward, rather than capturing the full range of possible outcomes over time—such as larger rewards after longer delays. Although the models can tell you if you've received a reward, they miss nuances, such as when and how much. After battling hunger—was the wait for the restaurant worth it?

An Unexpected Hint

Sousa and colleagues wondered if dopamine signaling is more complex than previously thought. Their new study was actually inspired by AI. An approach called distributional reinforcement learning estimates a range of possibilities and learns from trial and error rather than a single reward.

"What if different dopamine neurons were sensitive to distinct combinations of possible future reward features—for example, not just their magnitude, but also their timing?" said Sousa.

Harvard neuroscientists led by Naoshige Uchida had an answer. They recorded electrical activity from individual dopamine neurons in mice as the animals learned to lick up a water reward. At the beginning of each trial, the mice sniffed a different scent that predicted both the amount of water they might find—that is, the size of the reward—and how long until they might get it.

Each dopamine neuron had its own preference. Some were more impulsive and preferred immediate rewards, regardless of size. Others were more cautious, slowly ramping up activity that tracked reward over time. It's a bit like being extremely thirsty on a hike in the desert with limited water: Do you chug it all now, or ration it out and give yourself a longer runway?

Brain to AI

The brain recordings were like ensemble AI, where each model has its own viewpoint but the group collaborates to handle uncertainties.

The team also developed an algorithm, called time-magnitude reinforcement learning, or TMRL, that could plan future choices. Classic reinforcement-learning models only give out rewards at the end. This takes many cycles of learning before an algorithm homes in on the best decision. But TMRL rapidly maps a slew of choices, allowing humans and AI to pick the best ones with fewer cycles. The new model also includes internal states, like hunger levels, to further fine-tune decisions.

In one test, equipping algorithms with a dopamine-like "multidimensional map" boosted their performance in a simulated foraging task compared to standard reinforcement learning models.

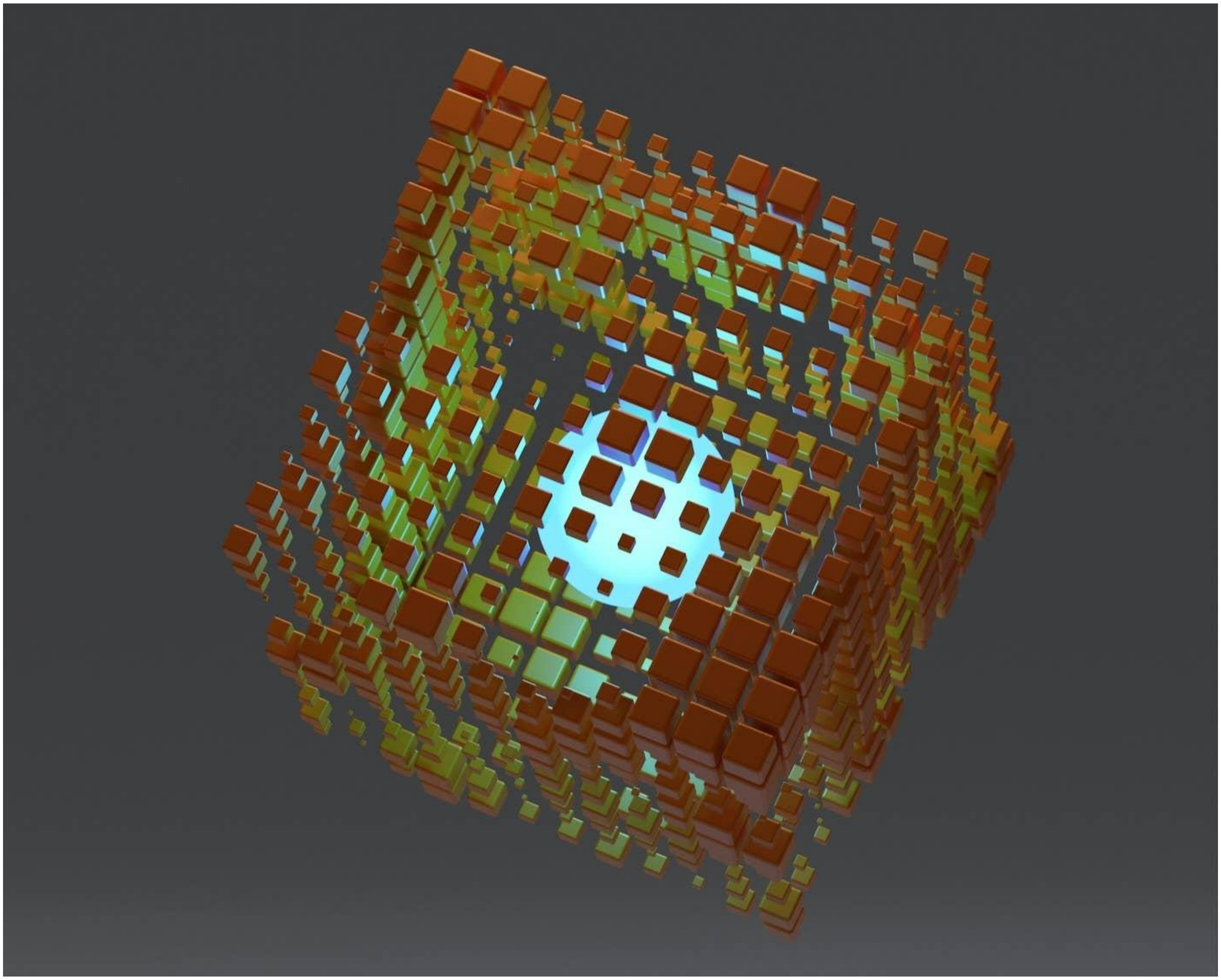
"Knowing in advance—at the start of an episode—the range and likelihood of rewards available and when they are likely to occur could be highly useful for planning and flexible behavior," especially in a complex environment and with different internal states, wrote Sousa and team.

The dual studies are the latest to showcase the power of AI and neuroscience collaboration. Models of the brain's inner workings can inspire more human-like AI. Meanwhile, AI is shining light into our own neural machinery, potentially leading to insights about neurological disorders.

Inspiration from the brain "could be key to developing machines that reason more like humans," said Paton.

Neurosymbolic AI Could Be the Answer to Hallucination in Large Language Models

Adding a dash of good 'ol fashioned AI to today's algorithms might bring about AI's third wave.



The main problem with big tech's experiment with artificial intelligence is not that it could take over humanity. It's that large language models (LLMs) like Open AI's ChatGPT, Google's Gemini, and Meta's Llama continue to get things wrong, and the problem is intractable. Known as hallucinations, the most prominent example was perhaps the case of US law professor Jonathan Turley, who was falsely accused of sexual harassment by ChatGPT in 2023.

OpenAI's solution seems to have been to basically "disappear" Turley by programming ChatGPT to say it can't respond to questions about him, which is clearly not a fair or satisfactory solution. Trying to solve hallucinations after the event and case by case is clearly not the way to go.

The same can be said of LLMs amplifying stereotypes or giving western-centric answers. There's also a total lack of accountability in the face of this widespread misinformation, since it's difficult to ascertain how the LLM reached this conclusion in the first place.

We saw a fierce debate about these problems after the 2023 release of GPT-4, the most recent major paradigm in OpenAI's LLM development. Arguably the debate has cooled since then, though without justification.

The EU passed its AI Act in record time in 2024, for instance, in a bid to be world leader in overseeing this field. But the act relies heavily on AI companies regulating themselves without really addressing the issues in question. It hasn't stopped tech companies from releasing LLMs worldwide to hundreds of millions of users and collecting their data without proper scrutiny.

Meanwhile, the latest tests indicate that even the most sophisticated LLMs remain unreliable. Despite this, the leading AI companies still resist taking responsibility for errors.

Unfortunately LLMs' tendencies to misinform and reproduce bias can't be solved with gradual improvements over time. And with the advent of agentic AI, where users will soon be able to assign projects to an LLM such as, say, booking their holiday or optimizing the payment of all their bills each month, the potential for trouble is set to multiply.

The emerging field of neurosymbolic AI could solve these issues, while also reducing the enormous amounts of data required for training LLMs. So what is neurosymbolic AI and how does it work?

The LLM Problem

LLMs work using a technique called deep learning, where they are given vast amounts of text data and use advanced statistics to infer patterns that determine what the next word or phrase in any given response should be. Each model—along with all the patterns it has learned—is stored in arrays of powerful computers in large data centers known as neural networks.

LLMs can appear to reason using a process called chain-of-thought, where they generate multi-step responses that mimic how humans might logically arrive at a conclusion, based on patterns seen in the training data.

Undoubtedly, LLMs are a great engineering achievement. They are impressive at summarizing text and translating and may improve the productivity of those diligent and knowledgeable enough to spot their mistakes. Nevertheless they have great potential to mislead because their conclusions are always based on probabilities—not understanding.

A popular workaround is called human-in-the-loop: making sure that humans using AIs still make the final decisions. However, apportioning blame to humans does not solve the problem. They'll still often be misled by misinformation.

LLMs now need so much training data to advance that we're having to feed them synthetic data, meaning data created by LLMs. This data can copy and amplify existing errors from its own source data, such that new models inherit the weaknesses of old ones. As a result, the cost of programming AI models to be more accurate after their training—known as post-hoc model alignment—is skyrocketing.

It also becomes increasingly difficult for programmers to see what's going wrong because the number of steps in the model's thought process becomes ever larger, making it harder and harder to correct for errors.

Neurosymbolic AI combines the predictive learning of neural networks with teaching the AI a series of formal rules that humans learn to be able to deliberate more reliably. These include logic rules, like "if a then b", which, for example, would help an algorithm learn that "if it's raining then everything outside is normally wet"; mathematical rules, like "if a = b and b = c then a = c"; and the agreed upon meanings of things like words, diagrams, and symbols. Some of these will be inputted directly into the AI system, while it will deduce others itself by analyzing its training data and performing "knowledge extraction."

The Third Wave

The first wave of AI in the 1980s, known as symbolic AI, was actually based on teaching computers formal rules that they could then apply to new information. Deep learning followed as the second wave in the 2010s, and many see neurosymbolic AI as the third.

It's easiest to apply neurosymbolic principles to AI in niche areas, because the rules can be clearly defined. So, it's no surprise that we've seen it first emerge in Google's AlphaFold, which predicts protein structures to help with drug discovery; and AlphaGeometry, which solves complex geometry problems.

For more broad-based AI models, China's DeepSeek uses a learning technique called "distillation" which is a step in the same direction. But to make neurosymbolic AI fully feasible for general models, there still needs to be more research to refine their ability to discern general rules and perform knowledge extraction.

It's unclear to what extent LLM makers are working on this already. They certainly sound like they're heading in the direction of trying to teach their models to think more cleverly, but they also seem wedded to the need to scale up with ever larger amounts of data.

The reality is that if AI is going to keep advancing, we will need systems that adapt to novelty from only a few examples, that check their understanding, that can multitask and reuse knowledge to improve data efficiency, and that can reason reliably in sophisticated ways.

This way, well-designed digital technology could potentially even offer an alternative to regulation, because the checks and balances would be built into the architecture and perhaps standardized across the industry. There's a long way to go, but at least there's a path ahead.

AlphaGenome: AI for better understanding the genome Published 25 JUNE 2025

How AlphaGenome works

Our AlphaGenome model takes a long DNA sequence as input — up to 1 million letters, also known as base-pairs — and predicts thousands of molecular properties characterising its regulatory activity. It can also score the effects of genetic variants or mutations by comparing predictions of mutated sequences with unmutated ones. Predicted properties include where genes start and where they end in different cell types and tissues, where they get spliced, the amount of RNA being produced, and also which DNA bases are accessible, close to one another, or bound by certain proteins. Training data was sourced from large public consortia including ENCODE, GTEx, 4D Nucleome and FANTOM5, which experimentally measured these properties covering important modalities of gene regulation across hundreds of human and mouse cell types and tissues.

Play video

Animation showing AlphaGenome taking one million DNA letters as input and predicting diverse molecular properties across different tissues and cell types.

The AlphaGenome architecture uses convolutional layers to initially detect short patterns in the genome sequence, transformers to communicate information across all positions in the sequence, and a final series of layers to turn the detected patterns into predictions for different modalities. During training, this computation is distributed across multiple interconnected Tensor Processing Units (TPUs) for a single sequence. This model builds on our previous genomics model, Enformer and is complementary to AlphaMissense, which specializes in categorizing the effects of variants within protein-coding regions. These regions cover 2% of the genome. The remaining 98%, called non-coding regions, are crucial for orchestrating gene activity and contain many variants linked to diseases. AlphaGenome offers a new perspective for interpreting these expansive sequences and the variants within them.

AlphaGenome's distinctive features

AlphaGenome offers several distinctive features compared to existing DNA sequence models:

Long sequence-context at high resolution

Our model analyzes up to 1 million DNA letters and makes predictions at the resolution of individual letters. Long sequence context is important for covering regions regulating genes from far away and base-resolution is important for capturing fine-grained biological details.

Previous models had to trade off sequence length and resolution, which limited the range of modalities they could jointly model and accurately predict. Our technical advances address this limitation without significantly increasing the training resources — training a single AlphaGenome model (without distillation) took four hours and required half of the compute budget used to train our original Enformer model.

Comprehensive multimodal prediction

By unlocking high resolution prediction for long input sequences, AlphaGenome can predict the most diverse range of modalities. In doing so, AlphaGenome provides scientists with more comprehensive information about the complex steps of gene regulation.

Efficient variant scoring

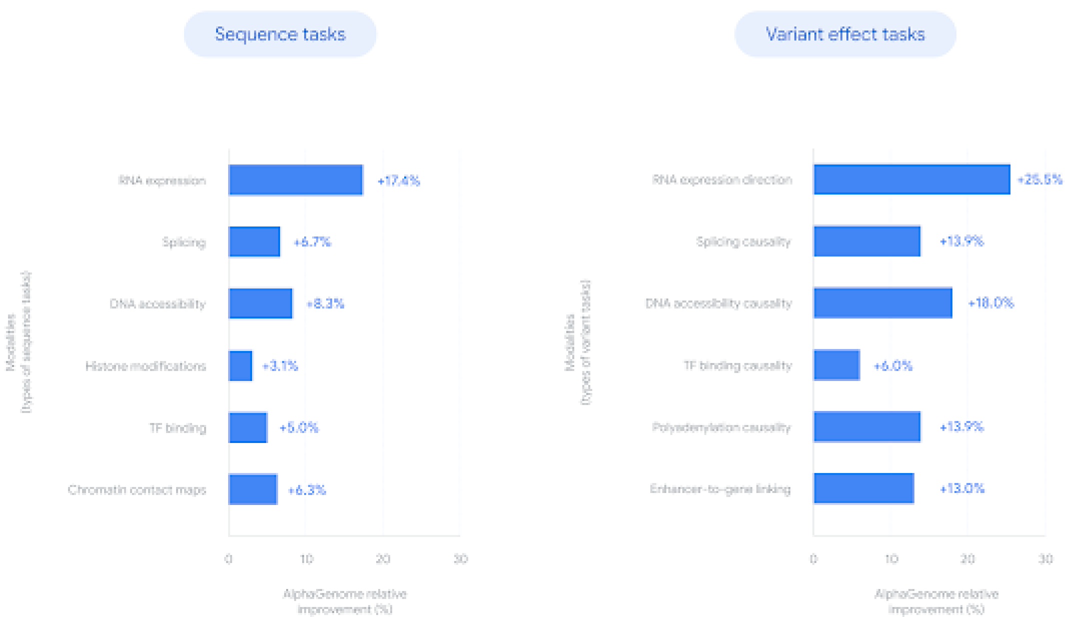
In addition to predicting a diverse range of molecular properties, AlphaGenome can efficiently score the impact of a genetic variant on all of these properties in a second. It does this by contrasting predictions of mutated sequences with unmutated ones, and efficiently summarising that contrast using different approaches for different modalities.

Novel splice-junction modeling

Many rare genetic diseases, such as spinal muscular atrophy and some forms of cystic fibrosis, can be caused by errors in RNA splicing — a process where parts of the RNA molecule are removed, or “spliced out”, and the remaining ends rejoined. For the first time, AlphaGenome can explicitly model the location and expression level of these junctions directly from sequence, offering deeper insights about the consequences of genetic variants on RNA splicing.

State-of-the-art performance across benchmarks

AlphaGenome achieves state-of-the-art performance across a wide range of genomic prediction benchmarks, such as predicting which parts of the DNA molecule will be in close proximity, whether a genetic variant will increase or decrease expression of a gene, or whether it will change the gene's splicing pattern.



Bar graph showing AlphaGenome's relative improvements on selected DNA sequence and variant effect tasks, compared against results for the current best methods in each category.

When producing predictions for single DNA sequences, AlphaGenome outperformed the best external models on 22 out of 24 evaluations. And when predicting the regulatory effect of a variant, it matched or exceeded the top-performing external models on 24 out of 26 evaluations.

This comparison included models specialized for individual tasks. AlphaGenome was the only model that could jointly predict all of the assessed modalities, highlighting its generality. Read more in our preprint.

The benefits of a unifying model

AlphaGenome's generality allows scientists to simultaneously explore a variant's impact on a number of modalities with a single API call. This means that scientists can generate and test hypotheses more rapidly, without having to use multiple models to investigate different modalities.

Moreover AlphaGenome's strong performance indicates it has learned a relatively general representation of DNA sequence in the context of gene regulation. This makes it a strong foundation for the wider community to build upon. Once the model is fully released, scientists will be able to adapt and fine-tune it on their own datasets to better tackle their unique research questions.

Finally, this approach provides a flexible and scalable architecture for the future. By extending the training data, AlphaGenome's capabilities could be extended to yield better performance, cover more species, or include additional modalities to make the model even more comprehensive.

It's a milestone for the field. For the first time, we have a single model that unifies long-range context, base-level precision and state-of-the-art performance across a whole spectrum of genomic tasks.
DR. CALEB LAREAU, MEMORIAL SLOAN KETTERING CANCER CENTER

A powerful research tool

AlphaGenome's predictive capabilities could help several research avenues:

1. Disease understanding: By more accurately predicting genetic disruptions, AlphaGenome could help researchers pinpoint the potential causes of disease more precisely, and better interpret the functional impact of variants linked to certain traits, potentially uncovering new therapeutic targets. We think the model is especially suitable for studying rare variants with potentially large effects, such as those causing rare Mendelian disorders.
2. Synthetic biology: Its predictions could be used to guide the design of synthetic DNA with specific regulatory function — for example, only activating a gene in nerve cells but not muscle cells.
3. Fundamental research: It could accelerate our understanding of the genome by assisting in mapping its crucial functional elements and defining their roles, identifying the most essential DNA instructions for regulating a specific cell type's function.

For example, we used AlphaGenome to investigate the potential mechanism of a cancer-associated mutation. In an existing study of patients with T-cell acute lymphoblastic leukemia (T-ALL), researchers observed mutations at particular locations in the genome. Using AlphaGenome, we predicted that the mutations would activate a nearby gene called TAL1 by introducing a MYB DNA binding motif, which replicated the known disease mechanism and highlighted AlphaGenome's ability to link specific non-coding variants to disease genes.

AlphaGenome will be a powerful tool for the field. Determining the relevance of different non-coding variants can be extremely challenging, particularly to do at scale. This tool will provide a crucial piece of the puzzle, allowing us to make better connections to understand diseases like cancer.
PROFESSOR MARC MANSOUR, UNIVERSITY COLLEGE LONDON

Current limitations

AlphaGenome marks a significant step forward, but it's important to acknowledge its current limitations.

Like other sequence-based models, accurately capturing the influence of very distant regulatory elements, like those over 100,000 DNA letters away, is still an ongoing challenge. Another priority for future work is further increasing the model's ability to capture cell- and tissue-specific patterns.

We haven't designed or validated AlphaGenome for personal genome prediction, a known challenge for AI models. Instead, we focused more on characterising the performance on individual genetic variants. And while AlphaGenome can predict molecular outcomes, it doesn't give the full picture of how genetic variations lead to complex traits or diseases. These often involve broader biological processes, like developmental and environmental factors, that are beyond the direct scope of our model.

We're continuing to improve our models and gathering feedback to help us address these gaps.

Enabling the community to unlock AlphaGenome's potential

AlphaGenome is now available for non-commercial use via our AlphaGenome API. Please note that our model's predictions are intended only for research use and haven't been designed or validated for direct clinical purposes.
Researchers worldwide are invited to get in touch with potential use-cases for AlphaGenome and to ask questions or share feedback through the community forum.

We hope AlphaGenome will be an important tool for better understanding the genome and we're committed to working alongside external experts across academia, industry, and government organizations to ensure AlphaGenome benefits as many people as possible.

Together with the collective efforts of the wider scientific community, we hope it will deepen our understanding of the complex cellular processes encoded in the DNA sequence and the effects of variants, and drive exciting new discoveries in genomics and healthcare.

JUNE 26, 2025

Engineers create first AI model specialized for chip design language

by NYU Tandon School of Engineering edited by Lisa Lock, reviewed by Andrew Zinin

From left to right: Jiang Hu, Editor in Chief of ACM TODAES; Ben Tan, paper co-author; Siddharth Garg, paper lead author; Helen Li, General Chair of Design Automation Conference. Credit: NYU Tandon School of Engineering

Researchers at NYU Tandon School of Engineering have created VeriGen, the first specialized artificial intelligence model successfully trained to generate Verilog code, the programming language that describes how a chip's circuitry functions.

The research just earned the ACM Transactions on Design Automation of Electronic Systems 2024 Best Paper Award, affirming it as a major advance in automating the creation of hardware description languages that have traditionally required deep technical expertise.

"General purpose AI models are not very good at generating Verilog code, because there's very little Verilog code on the Internet available for training," said lead author Institute Professor Siddharth Garg, who sits in NYU Tandon's Department of Electrical and Computer Engineering (ECE) and serves on the faculty of NYU WIRELESS and NYU Center for Cybersecurity (CCS). "These models tend to do well on programming languages that are well represented on GitHub, like C and Python, but tend to do a lot worse on poorly represented languages like Verilog."

Along with Garg, a team of NYU Tandon Ph.D. students, postdoctoral researchers, and faculty members Ramesh Karri and Brendan Dolan-Gavitt tackled this challenge by creating and distributing the largest AI training dataset of Verilog code ever assembled. They scoured GitHub to gather approximately 50,000 Verilog files from public repositories, and supplemented this with content from 70 Verilog textbooks. This data collection process required careful filtering and de-duplication to create a high-quality training corpus.

For their most powerful model, the researchers then fine-tuned Salesforce's open-source CodeGen-16B language model, which contains 16 billion parameters and was originally pre-trained on both natural language and programming code.

The computational demands were substantial. Training required three NVIDIA A100 GPUs working in parallel, with the model parameters alone consuming 30 GB of memory and the full training process requiring approximately 250 GB of GPU memory.

This fine-tuned model performed impressively in testing, outperforming commercial state-of-the-art models while being an order of magnitude smaller and fully open-source. In their evaluation, the fine-tuned CodeGen-16B achieved a 41.9% rate of functionally correct code versus 35.4% for the commercial code-davinci-002 model—with fine-tuning boosting accuracy from just 1.09% to 27%, demonstrating the significant advantage of domain-specific training.

"We've shown that by fine-tuning a model on that specific task you care about, you can get orders of magnitude reduction in the size of the model," Garg noted, highlighting how their approach improved both accuracy and efficiency. The smaller size enables the model to run on standard laptops rather than requiring specialized hardware.

The team evaluated VeriGen's capabilities across a range of increasingly complex hardware design tasks, from basic digital components to advanced finite state machines. While still not perfect—particularly on the most complex challenges—VeriGen demonstrated remarkable improvements over general-purpose models, especially in generating syntactically correct code.

The significance of this work has been recognized in the field, with subsequent research by NVIDIA in 2025 acknowledging VeriGen as one of the earliest and most important benchmarks for LLM-based Verilog generation, helping establish foundations for rapid advancements in AI-assisted hardware design.

The project's open-source nature has already sparked significant interest in the field. While VeriGen was the team's first model presented in the ACM paper, they've since developed an improved family of models called "CL Verilog" that perform even better.

These newer models have been provided to hardware companies including Qualcomm and NXP for evaluation of potential commercial applications. The work builds upon earlier NYU Tandon efforts including the 2020 DAVE (Deriving Automatically Verilog from English) project, advancing the field by creating a more comprehensive solution through large-scale fine-tuning of language models.

VeriGen complements other AI-assisted chip design initiatives from NYU Tandon aimed at democratizing hardware: their Chip Chat project created the first functional microchip designed through natural language conversations with GPT-4; Chips4All, supported by the National Science Foundation's (NSF's) Research Traineeship program, trains diverse STEM graduate students in chip design; and BASICS, funded through NSF's Experiential Learning for Emerging and Novel Technologies initiative, teaches chip design to non-STEM professionals.

More information: Shailja Thakur et al, VeriGen: A Large Language Model for Verilog Code Generation, *ACM Transactions on Design Automation of Electronic Systems* (2024). DOI: 10.1145/3643681
Provided by NYU Tandon School of Engineering

The Superintelligence Rush Is Here. This Is What Comes Next.

AGI? As in, artificial "general" intelligence? *That's so 2024.*

The AI industry just stopped pretending AGI is "someday" off in the future. Now, everyone's racing to build artificial "superintelligence" by 2027.

Just look at the marketing language:

Microsoft just claimed it's on the path to "medical superintelligence." Meta launched "Meta Superintelligence Labs" with Zuck's goal of "personal superintelligence for everyone."

Even OpenAI CEO Sam Altman says we're already in a "gentle singularity", which is basically a reference to superintelligence. Anthropic's CEO Dario Amodei predicts a "country of geniuses in a data center" by 2026-2027. *Superintelligence by another name, no?*

So what gives? Did we hit AGI and nobody told us? Where was the headline, fancy blog post, and podcast gauntlet to confirm it? Or is there actually something else going on entirely?

It's like the Turing test all over again. First, no one can hit it. Then, it's up for debate. Then, the debate just sorta… moved on?

First, the Race (Now - 2027): Clearly, whatever is happening right now is either…

1. The science is actually showing superintelligence is likely.
2. AGI has been achieved behind closed doors and they just haven't released it yet.
3. No one can agree, and therefore never will agree, on what AGI looks like (rebrand!)
4. Superintelligence is just a stickier (and therefore cooler) idea.

**There's also the theory that "ASI" is actually just a guy named Soham who works at 3-4 startups at one time (AGI = "A Genius Indian", ASI = "A Soham Indian" lol).*

Mark Zuckerberg proved Meta it is super serious about the idea of superintelligence by spending just shy of $15B on ScaleAI (a company central to training new AI), and by hiring 8 (and counting) top researchers from OpenAI, including Trapit Bansal (co-creator of the o1 reasoning model), Shengjia Zhao (co-creator of ChatGPT and GPT-4), and Shuchao Bi (co-creator of GPT-4o voice mode).

OpenAI, naturally, did not like this. Chief Research Officer Mark Chen sent an urgent internal memo that said, "I feel a visceral feeling right now, as if someone has broken into our home and stolen something." SamA called the move "not the craziest thing that would happen in OpenAI history" and said "missionaries will beat mercenaries." And of course, OpenAI's reassessing their researchers comp…

Put another way: *we're in the lawless, cutthroat season of the AI race, where companies believe they're approaching a winner-take-all moment and are willing to go all in to get there.*

So what happens if any of these cutthroats actually achieve superintelligence?

Vinod Khosla, an extremely prescient and legendary tech investor, offered this vision in a new interview with Jack Altman (*keep in mind, Vinod is an investor in OpenAI, but still*).

- **2025-2030: The AI Intern Era.** Those superintelligent systems become your coworkers. Every professional gets AI assistants smarter than Stanford grads. AI handles 80% of work in 80% of all jobs.
- **The 2030s: Corporate Extinction Event.** Fortune 500 companies die faster than ever. Someone builds a billion-dollar company with 10 employees. The superintelligent "interns" surpass their human bosses.
- **2040+: Work Becomes Optional.** You work because you want to, not because you need rent money. Humanoid robots handle physical labor. All expertise—medical, legal, educational—becomes free.

In sum, Khosla's 80% sure the real show ends with humanity never needing to work again.

His biggest fear? Not rogue AI. It's China using "good AI" (free healthcare, education) to export its politics globally. *But TBH, when you hear VCs talk* *about AI's benefits, they sound preeeetty communist themselves with all their chat of cheap goods and no jobs. Sam Lessin had a good meme where he said "*AI = communist, crypto = capitalist*."*

As for AGI… It's clear that we are in a Jagged AGI stage, where we have AI systems that are superhuman in some areas but fail at seemingly simple tasks

OpenAI's o3 exemplifies this perfectly: it achieved 87.5% on ARC-AGI (exceeding the 85% "AGI threshold"), and can create complete business plans from single prompts, yet consistently fails simple riddles that humans solve easily.

Similarly, Google's AlphaFold predicts protein structure "to within the width of an atom" but cannot (yet) apply this knowledge to other molecular problems.

This jagged nature explains why companies are simultaneously claiming AGI achievement while acknowledging significant limitations. Because we may already have AGI-level capabilities in many domains, superintelligence (where AI is smarter than humans) really is the next meaningful milestone.

Treats To Try.

1. *Guidde turns your screen recordings into professional video tutorials with AI-generated step-by-step narration and voiceover in 100+ languages.
2. **AI about to revamp your whole office suite:**

1. **Word:** Genspark's AI Docs turns your prompts into polished documents with automatic formatting and design (read more on X).
    1. **Alts:** Lex, which is a collaborative doc editor, Type, which is a solo writing app, or Sudowrite, which is for fiction-based writing.
2. **Excel:** Shortcut does your Excel work 10x faster than you can, with full file compatibility so you never need Excel again (here's a demo).
    1. **Alts:** AI Sheets from HuggingFace, Quadratic which is more for an open dev audience, or Julius which is more of an AI data analyst tool.
3. **Powerpoint:** Chronicle turns your bullet points into presentation slides with drag-and-drop widgets, or there's FlashDocs, which turns markdown into branded PowerPoint and Google Slides through their API (so you can make presentations with AI agents).
    1. **Alts:** Gamma, which is the #1 most popular presentation maker.
4. **Outlook:** Cora sorts your important emails and summarizes the rest into twice-daily briefs, or the recently acquired Superhuman, which makes you faster at processing email manually through lightning-fast keyboard shortcuts and a beautifully designed interface.
    1. **Alts:** Shortwave, a cheaper, popular alternative to Superhuman, or Thunderbird, which free and open-source.
5. **OH, and what about Edge / Chrome?** There's Dia, a browse that builds an assistant into every web page so you can write, learn, and shop smarter without switching tab.
    1. **Alts:** None that good atm, but Perplexity's Comet is on the way…here's a good article on the AI browser wars.

3. **Honorable mentions**, that aren't Office Suite replacements, but add-ons?
    1. Quillbot, which gives you writing tools anywhere via browser extension.

2. Plus AI, which is a chrome extension for G docs.
3. Equals, which is more of a specialized revenue dashboard for sales teams.
4. Formula Bot, which is a one-line prompt to formula/chart tool.

Try any or all of them out and see which ones work best for you!

- Perplexity launched a $200 a month plan for power users called Perplexity Max, which will give you unlimited Labs (which let you create dashboards, spreadsheets, presentations, web apps, you name it) and early access to things like Comet.
- Oracle's cloud business is booming, based on its deals with ByteDance (it will likely be part of the TikTok deal U.S. President Trump is negotiating) and OpenAI, who just signed on to rent another 4.5 gigawatts of servers.
- Microsoft confirmed plans to lay off 4% of its staff, which the company denies is part of an AI push, but some analysts have tied to their reduced margins from AI and $80B AI data center buildout.

JUNE 30, 2025

Brain-computer interface robotic hand control reaches new finger-level milestone

by Sara Pecchia, Carnegie Mellon University

edited by Sadie Harley, reviewed by Robert Egan

Experimental paradigm. Credit: *Nature Communications* (2025). DOI: 10.1038/s41467-025-61064-x

Robotic systems have the potential to greatly enhance daily living for the over one billion individuals worldwide who experience some form of disability. Brain-computer interfaces or BCIs present a compelling option by enabling direct communication between the brain and external devices, bypassing traditional muscle-based control.

While invasive BCIs have demonstrated the ability to control robotic systems with high precision, their reliance on risky surgical implantation and ongoing maintenance restricts their use to a limited group of individuals with severe medical conditions.

Carnegie Mellon University professor Bin He has spent over two decades investigating noninvasive BCI solutions, particularly those based on electroencephalography (EEG), that are surgery-free and adaptable across a range of environments.

His group has achieved a series of groundbreaking milestones using EEG-based BCIs, including the first successful flight of a drone, the first control of a robotic arm, and the first to control a robotic hand for continuous movement.

As detailed in a new study in *Nature Communications*, He's lab brings noninvasive EEG-based BCI one step closer to everyday use by demonstrating real-time brain decoding of individual finger movement intentions and control of a dexterous robotic hand at the finger level.

"Improving hand function is a top priority for both impaired and able-bodied individuals, as even small gains can meaningfully enhance ability and quality of life," explained Bin He, professor of biomedical engineering at Carnegie Mellon University.

"However, real-time decoding of dexterous individual finger movements using noninvasive brain signals has remained an elusive goal, largely due to the limited spatial resolution of EEG."

In a first-of-its-kind achievement for EEG-based BCI, He's group employed a real-time, noninvasive robotic control system that utilized movement execution and motor imagery of individual finger movements to drive corresponding robotic finger motions. Just by thinking about it, human subjects were able to successfully perform two- and three-finger control tasks.

This was accomplished with the assistance of a novel deep-learning decoding strategy and a network fine-tuning mechanism for continuous decoding from noninvasive EEG signals.

The goal moving forward is to build on this work to achieve more refined finger-level tasks, for instance, typing.

"The insights gained from this study hold immense potential to elevate the clinical relevance of noninvasive BCIs and enable applications across a broader population," He added.

"Our study highlights the transformative potential of EEG-based BCIs and their application beyond basic communication to intricate motor control."

More information: Yidan Ding et al, EEG-based brain-computer interface enables real-time robotic hand control at individual finger level, *Nature Communications* (2025). DOI: 10.1038/s41467-025-61064-x
Journal information: Nature Communications
Provided by Carnegie Mellon University

JUNE 30, 2025

The rise of 'artificial historians': AI as humanity's record-keeper

by Taylor & Francis

edited by Sadie Harley, reviewed by Andrew Zinin

Credit: AI-generated image

In documenting and recording society's collective data on an unprecedented scale, artificial intelligence is becoming humanity's historian—changing the way we record information for posterity.

But AI's inadvertent role as memory-keeper raises profound concerns for today's historians. Unlike human historians who explicitly document their methodologies, AI systems are creating the historical archives of the future without crucial transparency around how sources are selected, weighted, and interpreted.

This undermines a fundamental principle of historical scholarship, that methodologies should be visible and contestable. In the new book "Artificial Historians", historian Marnie Hughes-Warrington explores how AI systems are transforming historical records.

The author argues that AI is already deeply involved in history-making, generating 'most of the histories made around the globe' on a daily basis. Rather than seeing this as only a threat, the author encourages historians to see it as an opportunity to engage with AI development to ensure these systems reflect historical complexity.

History and nuanced understanding

Hughes-Warrington puts forward concerns around biases in data collection, specifically the "uneven and unfair collection of information about the past." When AI systems train on these biased historical records, they risk amplifying and perpetuating historical inequities, potentially cementing problematic narratives for future generations.

As well as this, some historical information may simply not be computable or readable by AI tools, giving an incomplete picture.

As well as the concerns around information gathering and transparency, Hughes-Warrington points out that AI misses the nuances of historical storytelling that humans inherently accept.

She explains that historical claims made by scholars and historians are never fully or perfectly true but are "partially grounded," meaning they refer to evidence outside themselves and invite testing. This complex understanding of historical information, or historical "truths," presents a challenge for AI systems trained to provide definitive answers.

When asked about world history topics, AI platforms tend to give similar, conventional responses that present a limited view of history, demonstrating that AI systems lack the nuanced understanding of historical context that human historians develop through years of study, Hughes-Warrington suggests.

"Information from the past might not be available or even computable, or presented in ways that make the use or combination of datasets difficult or even impossible," Hughes-Warrington explains.

"The contexts for data collection might also be ignored. If you knew that information was collected about people in financial or judicial distress, for instance, would you use it without thinking about their experiences? Most importantly, though, there may be an overconfidence in the development of algorithms or the detection of patterns."

AI is here to stay

"This hollowing out of history and its absorption into future, fiction, or geopolitics means that the historical expertise needed to make AI more effective and fair is missed. AI is not a threat to history if we see the invitation to be involved in its making," she explains.

"By bringing historical expertise to AI development, we can create more effective and fair artificial historians while preserving the critical thinking and contextual understanding that defines quality historical scholarship."

The text ultimately suggests that history-making is a complex, interpretive process that cannot be reduced to simple algorithms or rules. Hughes-Warrington challenges historians and AI technologists to think more deeply about how we define and create history.

"If history is the problem, then history is also the solution," Hughes-Warrington concludes.

More information: Marnie Hughes-Warrington. Artificial Historians. DOI: 10.4324/9781003275084
Provided by Taylor & Francis

JUNE 30, 2025

AI pinpoints promising materials that capture only CO_2 from air

by The Korea Advanced Institute of Science and Technology (KAIST)

edited by Lisa Lock, reviewed by Robert Egan

Concept diagram of direct air capture (DAC) technology and carbon capture using Metal-Organic Frameworks (MOFs). MOFs are promising porous materials capable of capturing carbon dioxide from the atmosphere, drawing attention as a core material for DAC technology. Credit: *Matter* (2025). DOI: 10.1016/j.matt.2025.102203
In order to help prevent the climate crisis, actively reducing already-emitted CO_2 is essential. Accordingly, direct air capture (DAC)—a technology that directly extracts only CO_2 from the air—is gaining attention. However, effectively capturing pure CO_2 is not easy due to water vapor (H_2O) present in the air.

KAIST researchers have successfully used AI-driven machine learning techniques to identify the most promising CO_2-capturing materials among metal-organic frameworks (MOFs), a key class of materials studied for this technology.

The research team, led by Professor Jihan Kim from the Department of Chemical and Biomolecular Engineering, in collaboration with a team at Imperial College London, published their research in the journal *Matter*.

The difficulty in discovering high-performance materials is due to the complexity of structures and the limitations of predicting intermolecular interactions. To overcome this, the research team developed a machine learning force field (MLFF) capable of precisely predicting the interactions between CO_2, water (H_2O), and MOFs. This new method enables calculations of MOF adsorption properties with quantum-mechanics-level accuracy at vastly faster speeds than before.

Using this system, the team screened more than 8,000 experimentally synthesized MOF structures, identifying more than 100 promising candidates for CO_2 capture. Notably, this included new candidates that had not been uncovered by traditional force-field-based simulations. The team also analyzed the relationships between MOF chemical structure and adsorption performance, proposing seven key chemical features that will help in designing new materials for DAC.

Concept diagram of adsorption simulation using Machine Learning Force Field (MLFF). The developed MLFF is applicable to various MOF structures and allows for precise calculation of adsorption properties by predicting interaction energies during repetitive Widom insertion simulations. It is characterized by simultaneously achieving high accuracy and low

computational cost compared to conventional classical force fields. Credit: *Matter* (2025). DOI: 10.1016/j.matt.2025.102203
This research is recognized as a significant advance in the DAC field, greatly enhancing materials design and simulation by precisely predicting MOF-CO_2 and MOF-H_2O interactions.

More information: Yunsung Lim et al, Accelerating CO_2 direct air capture screening for metal-organic frameworks with a transferable machine learning force field, *Matter* (2025). DOI: 10.1016/j.matt.2025.102203
Journal information: Matter
Provided by The Korea Advanced Institute of Science and Technology (KAIST)

JULY 2, 2025

How AI is improving accounting efficiency—without replacing jobs

by Seb Murray, Stanford University

edited by Lisa Lock, reviewed by Andrew Zinin

Credit: Pixabay/CC0 Public Domain
A new study finds that accountants are using AI to streamline workflows and close their books faster. The takeaway: AI isn't replacing accountants, but enhancing their work when combined with human expertise.

Accountants have long been seen as easy targets for automation. Accounting often tops lists of the most automatable jobs, owing to its reliance on routine tasks like data entry and reconciliation: exactly the kind of work computers—and now, artificial intelligence—are built to handle.

But a working paper by Jung Ho Choi, an assistant professor of accounting at Stanford Graduate School of Business, and Chloe Xie, Ph.D. '20, of MIT Sloan School of Management, challenges the idea that AI is here to simply replace accountants.

Instead, they find that accountants who use generative AI can support more clients, close the books faster, and provide higher-quality service. Rather than replacing bean counters and bookkeepers, AI helps them work more efficiently by automating repetitive tasks and flagging issues in real time, making it easier to complete reports quickly and accurately. And accountants themselves say that AI is making their jobs more manageable.

"AI helps with multitasking," Choi says. "In order to support each client, accountants have to pull information, connect bank transactions, track vendors—a lot of prework. AI assists with that setup, which means they can serve more clients, more efficiently."

The research draws on survey responses from 277 accountants and detailed task-level data for 79 small- and mid-sized firms that use AI-powered accounting tools. An analysis of both datasets shows that AI is taking over repetitive tasks, like transaction classification, allowing accountants to shift their attention to higher-value work.

According to the study, accountants who use AI support more clients per week and finalize monthly statements 7.5 days faster than those who use traditional methods. They also spend 8.5% less time on routine back-office processing. Instead of grinding through transactions line by line, they can redirect that time toward business communication, quality assurance, and client-facing advisory work.

Crucially, this increase in capacity does not appear to come at the expense of quality—if anything, standards improve. The study found that accounting firms using generative AI saw a 12% rise in reporting granularity, meaning they kept more detailed records. Instead of grouping expenses into broad categories like payroll, AI helped break them down into more specific categories like bonuses, benefits, or meals. This makes financial reports more informative and easier to analyze, audit, and act on.

"If you think about the early adoption of anything, there is generally some trade-off between quantity and quality," Xie says. "Whereas in this instance, perhaps surprisingly, the trade-off is not so sharp. That's probably most related to the fact that the technology is not here to replace the human being—it's here to augment the experts who are already in place."

Human expertise still matters

Not every accountant reaps the same performance gains from AI; it depends on their experience level—the more, the better. Senior accountants tend to treat AI as a collaborator. They are more discerning, stepping in when the system's confidence drops and applying human oversight where it's most needed.

Junior staff, on the other hand, are more likely to accept AI-generated outputs at face value, even when those outputs are flagged as uncertain. As a result, they see smaller performance gains. The study provides early evidence on potentially AI-generated errors flowing through human-in-the-loop accounting systems.

The takeaway isn't that AI is taking over accounting—it's that it works best when paired with human expertise. "Accountants are frequently the subject of a joke because so much of what they do is deeply routine and almost boring and procedural," Xie says. "But there's also a nontrivial proportion of an accountant's job that is very contextual and requires a lot of judgment."

Accountants are aware of these risks. In Choi and Xie's survey, 62% said they were concerned about AI-generated errors, and a significant number expressed anxiety about AI's effect on data security (43%) and job stability (37%).

Despite those concerns, nearly half of the accountants said generative AI tools helped them meet deadlines more reliably and improve accuracy. Almost two-thirds said that automating routine tasks was the single biggest benefit of adopting AI.

For now, the tools are mainly used for bookkeeping activities—the foundational, day-to-day tasks involved in recording and organizing a business's financial transactions. More complex areas like audit, tax strategy, and valuation are largely untouched. But that could change, according to Choi.

"We're hearing that AI can help synthesize information and standards for auditors to quickly get the gist of what matters," he says. "But that final judgment call—that's still a human decision."

More information: Human + AI in Accounting: Early Evidence from the Field. www.gsb.stanford.edu/faculty-r … early-evidence-field
Provided by Stanford University

JULY 2, 2025

AI might now be as good as humans at detecting emotion, political leaning and sarcasm in online conversations

by Ana Jovančević, The Conversation

edited by Lisa Lock, reviewed by Andrew Zinin

When we write something to another person, over email or perhaps on social media, we may not state things directly, but our words may instead convey a latent meaning—an underlying subtext. We also often hope that this meaning will come through to the reader.

But what happens if an artificial intelligence (AI) system is at the other end, rather than a person? Can AI, especially conversational AI, understand the latent meaning in our text? And if so, what does this mean for us?

Latent content analysis is an area of study concerned with uncovering the deeper meanings, sentiments and subtleties embedded in text. For example, this type of analysis can help us grasp political leanings present in communications that are perhaps not obvious to everyone.

Understanding how intense someone's emotions are or whether they're being sarcastic can be crucial in supporting a person's mental health, improving customer service, and even keeping people safe at a national level.

These are only some examples. We can imagine benefits in other areas of life, like social science research, policy-making and business. Given how important these tasks are—and how quickly conversational AI is improving—it's essential to explore what these technologies can (and can't) do in this regard.

Work on this issue is only just starting. Current work shows that ChatGPT has had limited success in detecting political leanings on news websites. Another study that focused on differences in sarcasm detection between different large language models—the technology behind AI chatbots such as ChatGPT—showed that some are better than others.

Finally, a study showed that LLMs can guess the emotional "valence" of words—the inherent positive or negative "feeling" associated with them. Our new study published in Scientific Reports tested whether conversational AI, inclusive of GPT-4—a relatively recent version of ChatGPT—can read between the lines of human-written texts.

The goal was to find out how well LLMs simulate understanding of sentiment, political leaning, emotional intensity and sarcasm—thus encompassing multiple latent meanings in one study. This study evaluated the reliability, consistency and quality of seven LLMs, including GPT-4, Gemini, Llama-3.1-70B and Mixtral 8 × 7B.

We found that these LLMs are about as good as humans at analyzing sentiment, political leaning, emotional intensity and sarcasm detection. The study involved 33 human subjects and assessed 100 curated items of text.

For spotting political leanings, GPT-4 was more consistent than humans. That matters in fields like journalism, political science, or public health, where inconsistent judgment can skew findings or miss patterns.

GPT-4 also proved capable of picking up on emotional intensity and especially valence. Whether a tweet was composed by someone who was mildly annoyed or deeply outraged, the AI could tell—although, someone still had to confirm if the AI was correct in its assessment. This was because AI tends to downplay emotions. Sarcasm remained a stumbling block both for humans and machines.

The study found no clear winner there—hence, using human raters doesn't help much with sarcasm detection.

Why does this matter? For one, AI like GPT-4 could dramatically cut the time and cost of analyzing large volumes of online content. Social scientists often spend months analyzing user-generated text to detect trends. GPT-4, on the other hand, opens the door to faster, more responsive research—especially important during crises, elections or public health emergencies.

Journalists and fact-checkers might also benefit. Tools powered by GPT-4 could help flag emotionally charged or politically slanted posts in real time, giving newsrooms a head start.

There are still concerns. Transparency, fairness and political leanings in AI remain issues. However, studies like this one suggest that when it comes to understanding language, machines are catching up to us fast—and may soon be valuable teammates rather than mere tools.

Although this work doesn't claim conversational AI can replace human raters completely, it does challenge the idea that machines are hopeless at detecting nuance.

Our study's findings do raise follow-up questions. If a user asks the same question of AI in multiple ways—perhaps by subtly rewording prompts, changing the order of information, or tweaking the amount of context provided—will the model's underlying judgements and ratings remain consistent?

Further research should include a systematic and rigorous analysis of how stable the models' outputs are. Ultimately, understanding and improving consistency is essential for deploying LLMs at scale, especially in high-stakes settings.

Provided by The Conversation

JULY 2, 2025

AI-powered assistive technologies are changing how we experience and imagine public space

by Ron Buliung, The Conversation

edited by Gaby Clark, reviewed by Andrew Zinin

New applications and the integration of artificial intelligence (AI) with wearable devices are changing the way users interact with their environments and each other. The impacts and reach of these new technologies have yet to be fully understood.

Connections between technologies and bodies is not a new thing for many disabled persons. Assistive technologies—tools and products designed to support people with disabilities—have played a part in mitigating built and institutional barriers experienced by disabled persons for decades.

While not strictly considered assistive, immersive and wearable technologies have the potential to change the relationship between disabled users and their experience of place.

For example, Ray-Ban's Meta glasses use AI to describe what the cameras are capturing using the Be My Eyes app. Using OpenAI's large language model, ChatGPT, this effectively turns a user's smartphone into a vision assistant.

Beyond wearables, some technologies are more closely tied to or integrated with the body. Examples include brain-computer interfaces, AI-enabled prosthetics and bone-anchored hearing aids.

The availability and production of environmental data from these technologies may impact how we relate to each other, how we move through and understand space, and how we engage with the physical environment around us at any given moment.

We're at a critical juncture where AI-enabled technologies used by individuals may profoundly impact our urban futures.

What happens, for example, when wearables make any "place" a digital work or play place? What does a largely private-sector, consumer-driven, AI-enabled digital intervention into a city's spaces mean for planning, zoning and taxation? What are the environmental costs of the global AI project?

And crucially, who gets to participate in this digital reimagining?

AI and the city

While access can be challenging—wearables are often costly—ableist thinking regarding the use of technology to render invisible Blind and/or Deaf people and culture is also a problem. Some people might naively assume that all Blind and Deaf people are universally seeking a bio-technological "miracle."

There are also other challenges: how a technology captures or describes its data may not match up to a user's pre-existing sense of place. Moreover, access to tech can produce some unintended consequences, including the erosion of in-person community building among disabled people.

Hearing loss of some kind affects around 1.5 billion people: I am one of those people. I am a disability studies scholar who wears behind-the-ear hearing aids to augment my hearing experience.

My hearing aids use AI and machine learning to sense and adjust my sound environment. They help me cope with the ways in which the places of my everyday life—such as my home or the lecture hall—are generally configured for people without hearing loss.

When I use my hearing aids, I find that the city has never sounded so wonderful, and yet sometimes irritatingly loud. The sound of birds is one thing; the grinding sound of a breaking subway is another entirely.

Cumulative exposure to noisy indoor and outdoor places of the city poses auditory health risks, such as noise-induced hearing loss or tinnitus, and can contribute to poor health more broadly. I have to be careful about ongoing noise exposure, and by adjusting the volume of my hearing aids, I can turn down the city when I want to.

Future bodies and urban futures

AI-powered technologies can exacerbate issues of access, privilege and freedom of movement. This happens both through who is able to purchase and use devices, as well as through data and their applications. Data may be biased in terms of race, gender, sexuality and disability.

Scientific research and media representations tend to highlight the benevolent possibilities of technologies for "repairing" bodies conceived as being functionally medically deficient.

Much less is said about disabled persons controlling the narrative, taking up key roles in the messy terrain of AI, machine learning and data governance, and in the planning and design of future cities.

Digital modeling

We are also witnessing growing interest in digital twinning—creating highly accurate digital models—of everything from human hearts to entire cities.

Whether rendered at the scale of the body or city, the motivation for twinning appears centered on planning and performance optimization—a quest for perfection. Like any model, we are dealing with an abstraction from reality. City twins seem to fail to capture many of the fine grain environmental barriers experienced by disabled persons.

Ownership limits

Not everyone can, should or wishes to be technologically "assisted" or augmented. There are medical, identity and culture, affordability, legal, moral and ethical concerns.

Other issues raised by brain-computer interface research, for example, include concerns about legal capacity and ownership of the self, including ownership of device-generated data.

In a study on the impact of neural technologies, researchers shared the legal repercussions relating to two disabled people deprived of voting rights in Spain. The person who recovered the ability to communicate autonomously using their finger and a computer had their rights restored, while the other, who used a human intermediary, did not.

Legal questions also arise regarding how liability is assigned when augmented bodies are injured or cause injuries to others.

Where does the person end and the technology begin, and vice versa? Who gets to decide?

Future technologies

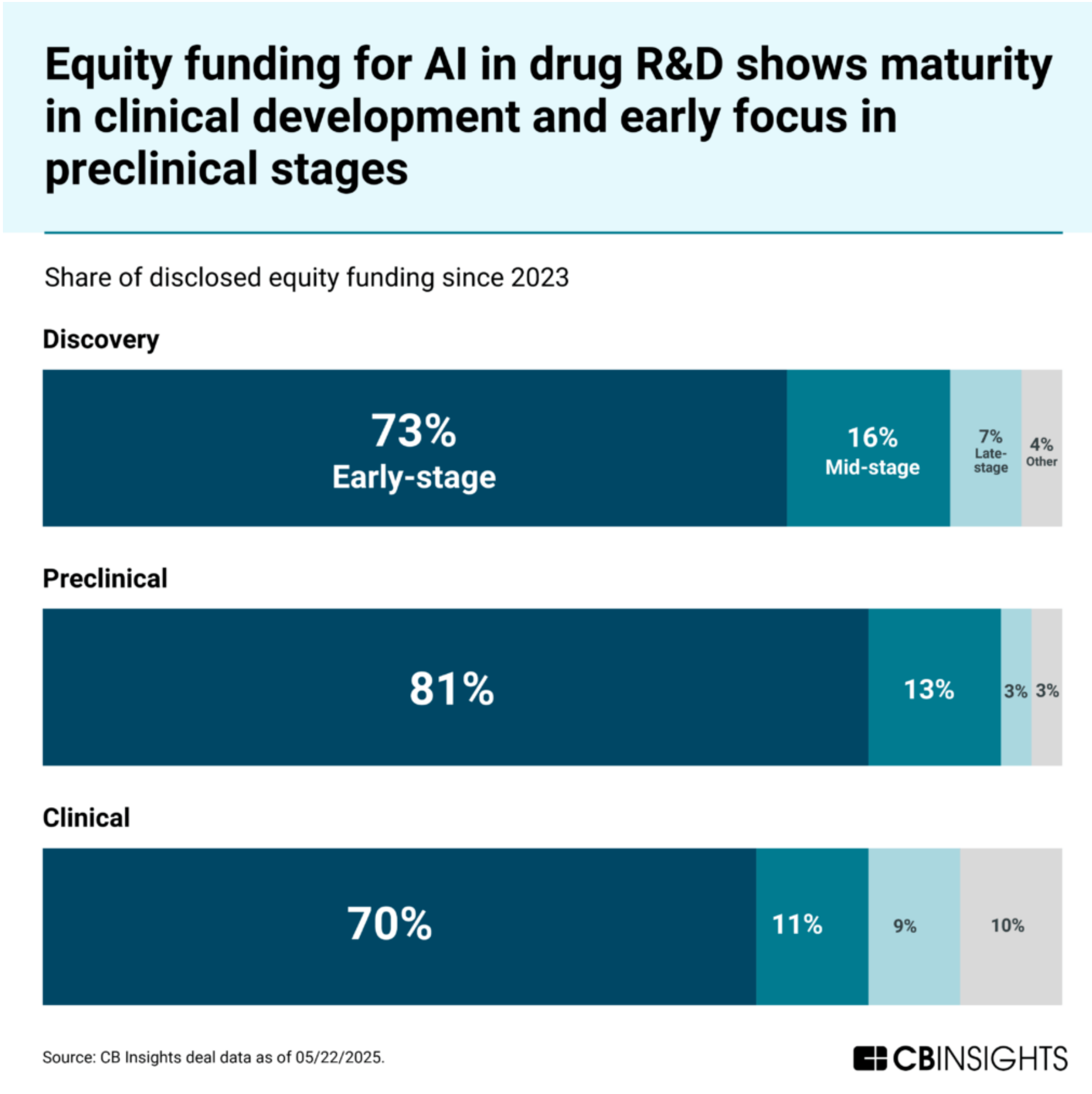
As the use of AI and assistive technologies increases in everyday urban life, we will need to address these questions sooner rather than later.

And if disabled persons are not adequately involved in these discussions and decisions, then cities will be less—rather than more—accessible.

Provided by The Conversation

METATRENDS

AI is Compressing Decades of Longevity Research into Weeks



What it is

AI is accelerating longevity research *millions-fold*.

Harvard lab is now **completing experiments in one month that would have previously taken "hundreds of thousands of years to accomplish"** using traditional methods. **AI is virtually screening**

trillions of molecules, identifying the precise combination needed to reverse aging, and **compressing decades of research into weeks**.

"AI is changing everything. What we do now in a month would've taken thousands of years to accomplish with traditional methods."
– David Sinclair, PhD

The pace of change is making even optimistic scientists' heads spin—and it's **fundamentally rewriting the timeline** for when age reversal becomes reality.

Why it matters

The Four-Lever Solution AI Just Cracked

For decades, scientists knew aging was complex, but they didn't know exactly *how* complex. Sinclair's team has now identified "the four main levers to reverse aging"—specific enzyme pathways that control the epigenome. The winning combination: "inhibit three of those and push one."

The challenge was finding a single molecule that could pull all four levers simultaneously. **“Five years ago, even if you asked pharmaceutical company with a billion dollars can you do that? They would’ve said:**

‘No way,’” explains Sinclair. "Doing one of those lever pulls is hard enough."

But AI changes everything.

Instead of chemists spending decades optimizing molecules by hand, Sinclair's lab can now "take all known molecules and virtually screen them in a couple of months against all four targets." So far, they have found a cocktail of three molecules that can activate the four main levers. Now they are looking to see if they can find a single molecule that could replace the entire cocktail.

The Power of AI Imaging Changes the Game

AI is also revolutionizing how rapidly Sinclair can evaluate the efficacy of an epigenetic reversing molecule. Sinclair's team developed an algorithm called "dash AI" that can accurately determine a cell’s age by imaging it under a microscope. "Within nanoseconds a computer can image and screen a skin cell and determine if it’s from a 20-year-old or a 93-year-old. So when we test an epigenetic age reversing molecule we can see if it works: if the age of the cell gets reduced, for example, from 93 years old down to 20 years old." This breakthrough enables ultra-rapid screening.

The Timeline Acceleration Is Stunning

Using AI and robotics, Sinclair's lab processes "trillions of molecules through this virtual screening process." They've identified "a hundred top candidates, some synthetic, some natural" that they can order from a chemical supply shop the same way you'd order something from Amazon.

Just five years ago, age reversal was purely theoretical. "In 2017, it was just a theory that we could reverse aging," Sinclair notes. By 2020, they proved it worked. Now, AI is accelerating the transition from the lab bench to human trials.

In success, these breakthrough compounds will reset aging in four weeks, and cost only hundreds of dollars for a four-week course of pills.

The Exponential Moment

We're witnessing something unprecedented: the convergence of AI and longevity science creating exponential progress in both fields.

"The pace of change is making my head spin off and I'm an optimist, but I just can't comprehend right now how fast things are going."
– David Sinclair, PhD

This is about compressing timelines for human benefit. Age reversal technologies that might have taken 30 years to develop are now racing toward human trials in *months*, not decades.