message to investors, and to the companies themselves, that there is a reputational damage that can result from this litigation," she says.

In an analysis of 120 climate cases, published on 17 April by the Grantham Research Institute, Setzer's team found that climate litigation can curb greenwashing in companies' advertisements — this includes making misleading statements about how climate-friendly certain products are, or disinformation about the effects of climate change (see go.nature.com/3unzqib). "With litigation being brought, companies are definitely communicating differently and being more cautious," she says.

**What's coming next in climate litigation?**

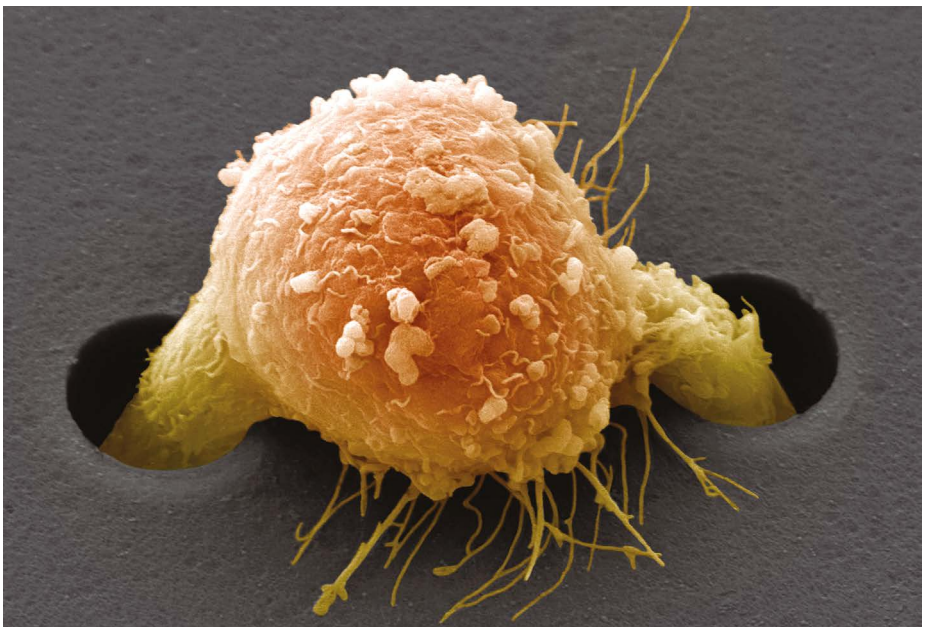
Maxwell thinks that people will bring more lawsuits that demand compensation from governments and companies for loss and damage caused by climate change. And more cases will be focused on climate adaptation — suing governments for not doing enough to prepare for and adjust to the effects of climate change, she says.

In an ongoing case from 2015, Peruvian farmer Saúl Luciano Lliuya argued that RWE, Germany's largest electricity producer, should contribute to the cost of protecting his hometown from floods caused by a melting glacier. He argued that planet-heating greenhouse gases emitted by RWE increase the risk of flooding.

More cases will be challenging an over-reliance by governments on carbon capture and storage (CCS) technologies — which remove carbon dioxide from the atmosphere and store it underground — in reaching emissions targets, says Maxwell. CCS technologies have not yet proved to work at a large scale. For instance, in February, researchers criticized the European Union for relying too much on CCS in its plans to cut greenhouse-gas emissions by 90% by 2040 compared with 1990 levels.

"There is a tendency now for companies and governments to say, we'll use carbon capture, we'll find some technology," says Setzer. "In the courts, we'll start seeing to what extent you can count on the future technologies, to what extent you really have to start acting now."

**By Carissa Wong**



STEVE GSCHMEISSNER/SPL

**A breast cancer cell (artificially coloured) climbs through a supportive film in the laboratory.**

# AI TRACES MYSTERIOUS METASTATIC CANCERS TO THEIR SOURCE

## Algorithm examines images of metastatic cells to identify the location of the primary tumour.

**By Smriti Mallapaty**

Some stealthy cancers remain undetected until they have spread from their source to distant organs. Now scientists have developed an artificial intelligence (AI) tool that outperforms pathologists at identifying the origins of metastatic cancer cells that circulate in the body. The proof-of-concept model could help doctors to improve the diagnosis of late-stage cancer and extend people's lives.

"That's a pretty significant finding — that it can be used as an assistive tool," says Faisal Mahmood, who studies AI applications in health care at Harvard Medical School in Boston, Massachusetts.

### Elusive origins

To treat metastatic cancers, doctors need to know where they came from. The origin of up to 5% of all tumours cannot be identified, and the prognosis for people whose primary cancer remains unknown is poor.

One method used to diagnose tricky metastatic cancers relies on tumour cells found in fluid extracted from the body. Clinicians examine images of the cells to work out which type of cancer cell they resemble. For example, breast cancer cells that migrate to the lungs still look like breast cancer cells.

Every year, of the 300,000 people with cancer who are newly treated at the hospital affiliated with Tianjin Medical University (TMU) in China, some 4,000 are diagnosed using such images, but around 300 people remain undiagnosed, says Tian Fei, a colorectal-cancer surgeon at TMU.

Tian, Li Xiangchun, a bioinformatics researcher who studies deep learning at TMU, and their colleagues wanted to develop a deep-learning algorithm to analyse these images and predict the origin of the cancers. Their results were published on 16 April (F. Tian *et al. Nature Med.* https://doi.org/mr2n; 2024).

### Tumour training

The researchers trained their AI model on some 30,000 images of cells found in abdominal or lung fluid from 21,000 people whose tumour of origin was known. They then tested their model on 27,000 images and found that there was an 83% chance that it would accurately predict the source of the tumour. Moreover, there was a 99% chance that the source of the tumour was included in the model's top three predictions.

Having a top-three list is useful because it can help clinicians to reduce the number

of extra — often intrusive — tests needed to identify a tumour's origins, says Mahmood. The predictions were restricted to 12 common sources of cancer, including the lungs, ovaries, breasts and stomach. Some other forms of cancer, including those originating in the prostate and kidneys, could not be identified, because they don't typically spread to fluid deposits in the abdomen and lungs, says Li.

When tested on some 500 images, the model was better than human pathologists at predicting a tumour's origin. This improvement was statistically significant.

The researchers also retrospectively assessed a subset of 391 study participants some four years after they had had cancer treatment. They found that those who had received treatment for the type of cancer that the model predicted were more likely to have survived, and lived longer, than were participants for whom the prediction did not match. "This is a pretty convincing argument" for using the AI model in a clinical setting, says Mahmood.

Mahmood has previously used AI to predict the origin of cancers from tissue samples (M. Y. Lu *et al. Nature* **594**, 106–110; 2021), and other teams have used genomic data. Combining the three data sources — cells, tissue and genomics — could further improve outcomes for people with metastatic cancers of unknown origins, he says.

# NEW BENCHMARKS NEEDED TO KEEP PACE WITH AI'S ADVANCE

## Stanford University's 2024 AI Index charts the meteoric rise of artificial-intelligence tools.

**By Nicola Jones**

Artificial intelligence (AI) systems, such as the chatbot ChatGPT, have become so advanced that they now very nearly match or exceed human performance in tasks including reading comprehension, image classification and competition-level mathematics, according to a report (see 'Speedy advances'). Rapid progress in the development of these systems also means that many common benchmarks and tests for assessing them are quickly becoming obsolete.

These are just a few of the headline findings from the Artificial Intelligence Index Report 2024, which was published on 15 April by the Institute for Human-Centered Artificial Intelligence at Stanford University in California (see go.nature.com/44ihnhx). The report charts the meteoric progress in machine-learning systems over the past decade.

In particular, the report says, new ways of assessing AI — for example, evaluating their performance on complex tasks, such as reasoning — are becoming more and more necessary. "A decade ago, benchmarks would serve the community for five to ten years", whereas now they often become irrelevant in just a few years, says Nestor Maslej, a social scientist at Stanford and editor-in-chief of the AI Index. "The pace of gain has been startlingly rapid."

Stanford's annual AI Index, first published in 2017, is compiled by a group of academic and industry specialists to assess the field's technical capabilities, costs, ethics and more — with an eye to informing researchers, policymakers and the public. This year's report, which is more than 400 pages long and was copy-edited and tightened with the aid of AI tools, notes that AI-related regulation in the United States is sharply rising. But the lack of standardized assessments for responsible use of AI makes it difficult to compare systems in terms of the risks that they pose.

The rise in the use of AI in science is also highlighted in this year's edition: for the first time, it dedicates an entire chapter to scientific applications, highlighting projects including Graph Networks for Materials Exploration (GNoME), a project from Google DeepMind that aims to help chemists discover materials, and GraphCast, another DeepMind tool, which does rapid weather forecasting.

### Growing up

The current AI boom — built on neural networks and machine-learning algorithms — dates back to the early 2010s. The field has since rapidly expanded. For example, the number of AI coding projects on GitHub, a common platform for sharing code, increased from about 800 in 2011 to 1.8 million last year. And journal publications about AI roughly tripled over this period, the report says.

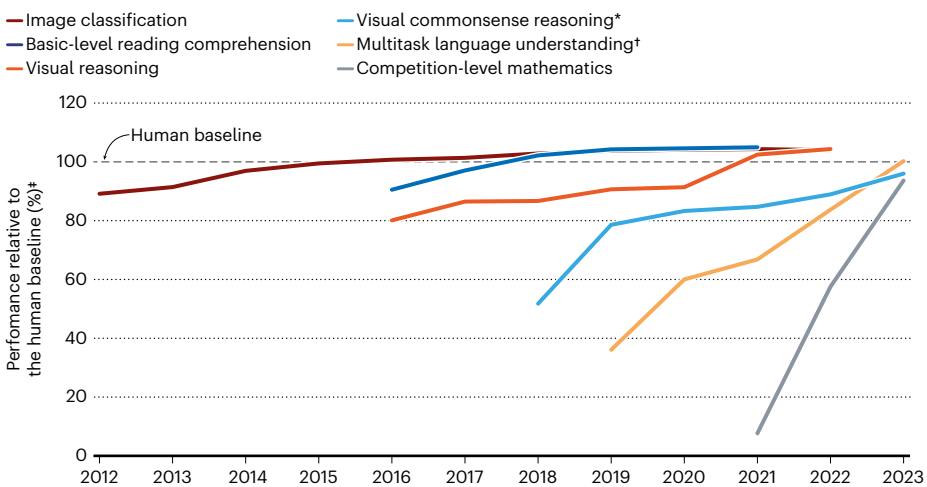
Much of the cutting-edge work on AI is being done in industry: that sector produced 51 notable machine-learning systems last year, whereas academic researchers contributed 15. "Academic work is shifting to analysing the models coming out of companies — doing a deeper dive into their weaknesses," says Raymond Mooney, director of the AI Lab at the University of Texas at Austin, who wasn't involved in the report.

That includes developing tougher tests to assess the visual, mathematical and even moral-reasoning capabilities of large language models (LLMs), which power chatbots. One of the latest tests is the Graduate-Level Google-Proof Q&A Benchmark (GPQA), developed last year by a team including machine-learning researcher David Rein at New York University (D. Rein *et al.* Preprint at arXiv https://doi.org/mr2k; 2023).

The GPQA, consisting of more than 400 multiple-choice questions, is tough: PhD-level scholars could correctly answer

### SPEEDY ADVANCES

In the past several years, some AI systems have surpassed human performance on certain benchmark tests, and others have made rapid progress.




*Requires an AI system to answer questions about an image and provide a rationale for why its answers are true. †Tests an AI model's knowledge and problem-solving ability with regard to 57 subjects, including broader topics such as mathematics and history, and narrower areas such as law and ethics. ‡Data indicate the best performance of an AI model that year.

SOURCE: ARTIFICIAL INTELLIGENCE INDEX REPORT 2024.