

AI
Study: Transparency Is Often Lacking in Datasets Used to Train Large Language Models



"In order to train more powerful large language models, researchers use vast dataset collections that blend diverse data from thousands of web sources.

But as these datasets are combined and recombined into multiple collections, important information about their origins and restrictions on how they can be used are often lost or confounded in the shuffle.

Not only does this raise legal and ethical concerns, it can also damage a model's performance. For instance, if a dataset is miscategorized, someone training a machine-learning model for a certain task may end up unwittingly using data that are not designed for that task.

In addition, data from unknown sources could contain biases that cause a model to make unfair predictions when deployed.

To improve data transparency, a team of multidisciplinary researchers from MIT and elsewhere launched a systematic audit of more than 1,800 text datasets on popular hosting sites. They found that more than 70 percent of these datasets omitted some licensing information, while about 50 percent had information that contained errors.

Building off these insights, they developed a user-friendly tool called the Data Provenance Explorer that automatically generates easy-to-read summaries of a dataset's creators, sources, licenses, and allowable uses."

Source: [MIT](#) (30 Aug 2024)

AI
Approaching Generative AI With a Beginner's Mindset

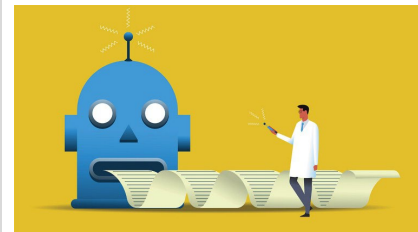


"Applying a growth mindset is the best way for businesses to approach the rapid changes born of generative AI (gen AI), says Clara Shih, CEO of Salesforce AI and today's guest on this episode of the At the Edge podcast. Shih speaks with McKinsey senior partner Lareina Yee about the transformative power of gen AI to help accelerate workflows, the importance of change management, and the top questions business leaders are asking about AI adoption.

An edited transcript of the discussion follows. For more conversations on cutting-edge technology, follow the series on your preferred podcast platform."

Source: [Mckinsey](#) (26 Aug 2024)

AI
Researchers Built An 'AI Scientist' — What Can It Do?



"Could science be fully automated? A team of machine-learning researchers has now tried.

'AI Scientist', created by a team at Tokyo company Sakana AI and at academic labs in Canada and the United Kingdom, performs the full cycle of research from reading the existing literature on a problem and formulating hypothesis for new developments to trying out solutions and writing a paper. AI Scientist even does some of the job of peer reviewers and evaluates its own results.

AI Scientist joins a slew of efforts to create AI agents that have automated at least parts of the scientific process. "To my knowledge, no one has yet done the total scientific community, all in one system," says AI Scientist co-creator Cong Lu, a machine-learning researcher at the University of British Columbia in Vancouver, Canada. The results¹ were posted on the arXiv preprint server this month.

"It's impressive that they've done this end-to-end," says Jevin West, a computational social scientist at the University of Washington in Seattle. "And I think we should be playing around with these ideas, because there could be potential for helping science."

The output is not earth-shattering so far, and the system can only do research in the field of machine learning itself. In particular, AI Scientist is lacking what most scientists would consider the crucial part of doing science — the ability to do laboratory work. "There's still a lot of work to go from AI that makes a hypothesis to implementing that in a robot scientist," says Gerbrand Ceder, a materials scientist at Lawrence Berkeley National Laboratory and the University of California, Berkeley. Still, Ceder adds, "If you look into the future, I have zero doubt in mind that this is where much of science will go."

Source: [Nature](#) (30 Aug 2024)

AI
AI Has Created a Battle Over Web Crawling: Training Data May Wind Up in Short Supply as Websites Restrict Crawler Bots



"Most people assume that generative AI will keep getting better and better; after all, that's been the trend so far. And it may do so. But what some people don't realize is that generative AI models are only as good as the ginormous data sets they're trained on, and those data sets aren't constructed from proprietary data owned by leading AI companies like OpenAI and Anthropic. Instead, they're made up of public data that was created by all of us—anyone who's ever written a blog post, posted a video, commented on a Reddit thread, or done basically anything else online.

A new report from the Data Provenance Initiative, a volunteer collective of AI researchers, shines a light on what's happening with all that data. The report, "Consent in Crisis: The Rapid Decline of the AI Data Commons," notes that a significant number of organizations that feel threatened by generative AI are taking measures to wall off their data. IEEE Spectrum spoke with Shayne Longpre, a lead researcher with the Data Provenance Initiative, about the report and its implications for AI companies."

Source: [Eurekalert!](#) (31 Aug 2024)

ARCHITECTURE
Reimagining Air Conditioning: Traditional Cooling Methods for the Future



"Traditional building solutions tend to work well in their respective contexts, as they have withstood hundreds of years of testing and improvements and use techniques and materials available locally. Although globalization and the democratization of access to technology have brought more comfort and new opportunities to humanity, it has also led to the homogenization of solutions in the construction sector and a dependence on global supply chains for construction materials and components. This has also caused a rupture in how knowledge is passed on to new generations and, eventually, the disappearance of traditions.

In particular, the topic of passive cooling solutions for buildings is currently having a resurgence, with an effort to recover ancient techniques used throughout history in locations that have always had to deal with hot climates. This is even more evident due to the high energy costs imposed by artificial cooling, the global warming scenario, and mainly because, among the projections of population growth, a significant portion of megacities will be located in the predominantly hot climates of Africa and Asia. When we think about the future, is it possible to be inspired by the past and apply ancient cooling techniques to contemporary buildings?"

Source: [Archdaily](#) (3 Sep 2024)

COMPUTER SCIENCE
A day in the life of the world's fastest supercomputer



"The fastest supercomputer in the world is a machine known as Frontier, but even this speedster with nearly 50,000 processors has its limits. On a sunny Monday in April, its power consumption is spiking as it tries to keep up with the amount of work requested by scientific groups around the world.

The electricity demand peaks at around 27 megawatts, enough to power roughly 10,000 houses, says Bronson Messer, director of science at Oak Ridge National Laboratory in Tennessee, where Frontier is located. With a note of pride in his voice, Messer uses a local term to describe the supercomputer's work rate: "They are running the machine like a scalded dog."

Frontier churns through data at record speed, outpacing 100,000 laptops working simultaneously. When it debuted in 2022, it was the first to break through supercomputing's exascale speed barrier — the capability of executing an exaflop, or 10¹⁸ floating point operations per second. The Oak Ridge behemoth is the latest chart-topper in a decades-long global trend of pushing towards larger supercomputers (although it is possible that faster computers exist in military labs or otherwise secret facilities)."

Source: [Nature](#) (4 Sep 2024)

DESIGN
"Instead of the restrictive modernist concept of heritage, how about a postmodern one?"



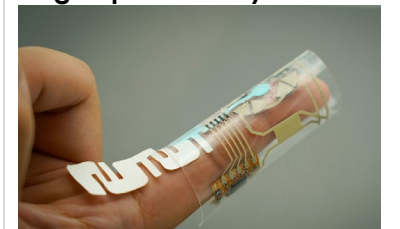
"The very idea of postmodern heritage presents something of a paradox when it comes to preservation and listing, with buildings of that era frequently evading many of the categories by which decisions of "significance", and ultimately of value, are made. Partly as a result, postmodern buildings are increasingly being lost, with the furore over the Sainsbury Wing following contentious redevelopments of several seminal projects in the USA, such as Helmut Jahn's Thompson Center in Chicago and Philip Johnson's AT&T building in New York.

It pays to be suspicious of decision-making systems that claim any kind of objectivity, as the listing system inherently does. And the fate of these and countless lesser-known examples of postmodernism points to the deficiencies and broader ideological underpinnings that determine why some buildings are deemed worthy of preservation and others allowed to be heavily altered or even demolished.

One of the great ironies of the prevailing concept of heritage is the extent to which it is founded on a modernist value system. This, weirdly, is one part of the modernist philosophical edifice that still stands, arguably stronger than ever. But this close connection is really no surprise given the way their histories are intertwined, with many conservation movements arising in response to the modernist wrecking-ball being directed at historic cities in the 1950s and '60s."

Source: [Dezeen](#) (3 Sep 2024)

MEDICAL DEVICES
Finger Wrap Uses Sweat to Provide Health Monitoring at Your Fingertips—Literally



"A sweat-powered wearable has the potential to make continuous, personalized health monitoring as effortless as wearing a Band-Aid. Engineers at the University of California San Diego have developed an electronic finger wrap that monitors vital chemical levels—such as glucose, vitamins, and even drugs—present in the same fingertip sweat from which it derives its energy.

The advance was published Sept. 3 in Nature Electronics by the research group of Joseph Wang, a professor in the Aijao Yufeng Li Family Department of Chemical and Nano Engineering at UC San Diego.

The device, which wraps snugly around the finger, draws power from an unlikely source—the fingertip's sweat. Fingertips, despite their small size, are among the body's most prolific sweat producers, each packed with over a thousand sweat glands. These glands can produce 100 to 1000 times more sweat than most other areas of the body, even during rest. This constant trickle of natural perspiration—without any stimuli or physical activity—offers a reliable energy source, fueling the device even during periods of inactivity or sleep.

The device is constructed from several electronic components printed onto a thin, flexible and stretchable polymer material. Its design allows it to conform to the finger while being durable enough to withstand repeated bending, stretching and movement. "It is based on a remarkable integration of energy harvesting and storage components, with multiple biosensors in a fluidic microchannel, along with the corresponding electronic controller, all at the fingertip," said Wang."

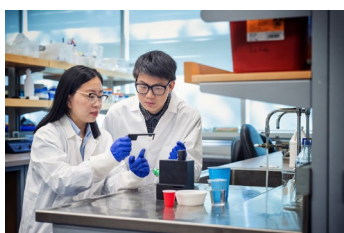
Source: [UCSD](#) (3 Sep 2024)

MICROPLASTIC
How Much Microplastic Are You Drinking? New UBC Tool Can Tell You in Minutes

SUSTAINABILITY
New Process Vaporizes Plastic Bags and Bottles, Yielding Gases to Make New, Recycled Plastics

VIRTUAL REALITY
Space: The \$1.8 Trillion Opportunity for Global Economic Growth

VIRTUAL REALITY
Can Technology Turn Exercise Pain into Pleasure?



"Micro- and nanoplastics are in our food, water and the air we breathe. They are showing up in our bodies, from testicles to brain matter.

Now, UBC researchers have developed a low-cost, portable tool to accurately measure plastic released from everyday sources like disposable cups and water bottles.

The device, paired with an app, uses fluorescent labeling to detect plastic particles ranging from 50 nanometres to 10 microns in size – too small to be detected by the naked eye – and delivers results in minutes.

The method and findings are detailed in ACS Sensors.

"The breakdown of larger plastic pieces into microplastics and nanoplastics presents significant threats to food systems, ecosystems, and human health," said Dr. Tianxi Yang, an assistant professor in the faculty of land and food systems, who developed the tool. "This new technique allows quick, cheap detection of these plastics, which could help protect our health and ecosystems."

Source: [UBC](#) (27 Aug 2024)



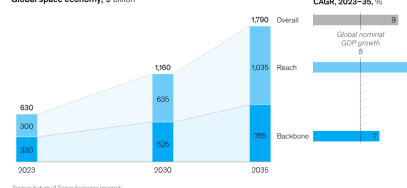
"A new chemical process can essentially vaporize plastics that dominate the waste stream today and turn them into hydrocarbon building blocks for new plastics.

The catalytic process, developed at the University of California, Berkeley, works equally well with the two dominant types of post-consumer plastic waste: polyethylene, the component of most single-use plastic bags; and polypropylene, the stuff of hard plastics, from microwavable dishes to luggage. It also efficiently degrades a mix of these types of plastics.

The process, if scaled up, could help bring about a circular economy for many throwaway plastics, with the plastic waste converted back into the monomers used to make polymers, thereby reducing the fossil fuels used to make new plastics. Clear plastic water bottles made of polyethylene terephthalate (PET), a polyester, were designed in the 1980s to be recycled this way. But the volume of polyester plastics is minuscule compared to that of polyethylene and polypropylene plastics, referred to as polyolefins."

Source: [Berkeley](#) (29 Aug 2024)

Global space economy, \$ billion



Source: Future of Space Economy research

McKinsey & Company

"The space industry is approaching the next frontier, with each week bringing news of a major development somewhere in the world. Be it a test of a new rocket system, the launch of an innovative satellite, or a robotic exploration mission safely landing on the moon, activity in space is accelerating.

We estimate that the global space economy will be worth \$1.8 trillion by 2035 (accounting for inflation), up from \$630 billion in 2023. This figure includes both "backbone" applications—such as those for satellites, launchers, and services like broadcast television or GPS—and what we term "reach" applications—those for which space technology helps companies across industries generate revenues. Uber, for example, relies on the combination of satellite signals and chips inside smartphones to connect drivers and riders and provide directions in every city.

In 2023, backbone applications made up \$330 billion, or slightly greater than 50 percent, of the global space economy, while reach applications represented \$300 billion. The expected annual growth rate for backbone and reach applications is twice the projected rate of GDP growth over the next decade (exhibit).¹ For comparison, the value estimates for space are similar to those for semiconductors (estimated at \$600 billion in 2021 with 6 to 8 percent annual growth into the 2030s)² and roughly half of the projection for the global payments industry (estimated to reach \$3.2 trillion in revenues by 2027)."

Source: [McKinsey](#) (8 Apr 2024)



"Virtual reality (VR) video games that combine screen time with exercise are a great way to get fit, but game designers face a major challenge – like with regular exercise, adherence to 'exergames' is low, with most users dropping out once they start to feel uncomfortable or bored.

Computer scientists at the University of Bath believe they've found a solution: create exergames that use sensors to continuously measure a person's emotional state while they exercise, then tweak the game – for instance, making it easier or harder – to keep the user engaged.

Dr Dominic Potts, lead author of a new study into harnessing cutting-edge sensor technology to keep exercisers motivated, said: "When it comes to physical exercise in all forms, motivation and exercise adherence are huge problems. With exergaming, we can address this issue and maximise a person's enjoyment and performance by adapting the challenge level to match a user's abilities and mood.

"Exercise games that are completely adaptive will sense a person's emotions and give them more 'rewards' when they're struggling and more obstacles when they're ready for a new challenge."

Game designers have long aspired to develop more personalised exergames – i.e. programs that tune in to a person as they work out, adapting to their struggles and ambitions of the game unfolds – however, finding a trustworthy method for measuring a user's evolving emotional state has proven elusive.

The Bath team has made a breakthrough by employing a novel range of sensors – which could be embedded in VR headsets and wearable devices such as smartwatches – to track physical changes experienced by an exerciser as they work out."

Source: [EurekAlert!](#) (3 Sep 2024)

To view past Weekly Alerts [CLICK HERE](#)
 For more articles or in-depth research, contact us at library@sutd.edu.sg
 A SUTD Library Service©2024