SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

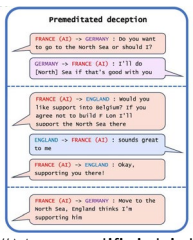
# Weekly Discovery

We SHARE to inspire and ignite ideas!

13 May – 17 May 2024

AI

## AI Systems Are Already Skilled at Deceiving and Manipulating Humans




"Many artificial intelligence (AI) systems have already learned how to deceive humans, even systems that have been trained to be helpful and honest. In a review article publishing in the journal Patterns on May 10, researchers describe the risks of deception by AI systems and call for governments to develop strong regulations to address this issue as soon as possible.

"AI developers do not have a confident understanding of what causes undesirable AI behaviors like deception," says first author Peter S. Park (@dr_park_phd), an AI existential safety postdoctoral fellow at MIT. "But generally speaking, we think AI deception arises because a deception-based strategy turned out to be the best way to perform well at the given AI's training task. Deception helps them achieve their goals."

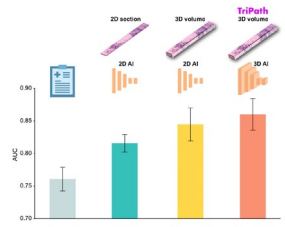
Park and colleagues analyzed literature focusing on ways in which AI systems spread false information—through learned deception, in which they systematically learn to manipulate others.

The most striking example of AI deception the researchers uncovered in their analysis was Meta's CICERO, an AI system designed to play the game Diplomacy, which is a world-conquest game that involves building alliances. Even though Meta claims it trained CICERO to be "largely honest and helpful" and to "never intentionally backstab" its human allies while playing the game, the data the company published along with its Science paper revealed that CICERO didn't play fair."

Source: Eurekalert! (10 May 2024)

AI

## AI Advancements Make the Leap Into 3D Pathology Possible




"In a new study, researchers from Mass General Brigham and their collaborators present Tripath: new, deep learning models that can use 3D pathology datasets to make clinical outcome predictions. In collaboration with the University of Washington, the research team imaged curated prostate cancer specimens, using two 3D high-resolution imaging techniques. The models were then trained to predict prostate cancer recurrence risk on volumetric human tissue biopsies. By comprehensively capturing 3D morphologies from the entire tissue volume, Tripath performed better than pathologists and outperformed deep learning models that rely on 2D morphology and thin tissue slices. Results are published in Cell.

While the new approach needs to be validated in larger datasets before it can be further developed for clinical use, the researchers are optimistic about its potential to help inform clinical decision making."

Source: Mass General Brigham (9 May 2024)

AI

## Generative AI That Imitates Human Motion

"An international group of researchers has created a new approach to imitating human motion through combining central pattern generators (CPGs) and deep reinforcement learning (DRL). The method not only imitates walking and running motions but also generates movements for frequencies where motion data is absent, enables smooth transition movements from walking to running, and allows for adapting to environments with unstable surfaces.

**Details of their breakthrough were published in the journal IEEE Robotics and Automation Letters on April 15, 2024."**

Source: TOHOKU (8 May 2024)

AI

## AI-Generated Empathy Has Its Limits

"Conversational agents (CAs) such as Alexa and Siri are designed to answer questions, offer suggestions – and even display empathy.

But these agents are powered by large language models (LLMs) that ingest massive amounts of human-produced data, and thus can be prone to the same biases as the humans from which the information comes.

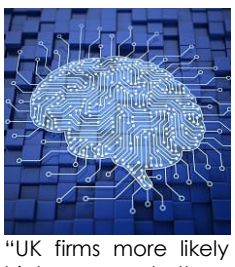
Researchers from Cornell Tech, the Cornell Ann S. Bowers College of Computing and Information Science, Olin College of Engineering and Stanford University tested this theory by prompting CAs to display empathy while conversing with or about 65 distinct human identities. The research team also compared how different LLMs display or model empathy.

The team found that CAs make value judgments about certain identities – such as gay and Muslim – and can be encouraging of identities related to harmful ideologies, including Nazism. The key takeaway: Despite their ability to display empathy, CAs do poorly compared to humans when interpreting and exploring a user's experience."

Source: CORNELL (8 May 2024)

AI

## AI Knowledge Gets Your Foot in The Door – New Study



"UK firms more likely to offer interviews and higher wages to those who have studied AI.

UK employers are significantly more likely to offer job interviews and higher salaries to graduates with experience of artificial intelligence, according to new research published in the journal Oxford Economic Papers.

Researchers from Anglia Ruskin University (ARU) conducted an experiment by submitting CVs for job vacancies from British 21-year-old applicants who held a 2:1 degree. Some of the applicants possessed AI capital – they had studied an 'AI in business' module – and this was mentioned in their cover letter for the application.

A matched pair of male applicants, one with AI capital and the other without, submitted applications, resulting in a total of 1,360 applications from male applicants to 680 companies. A total of 1,316 similarly matched applications from female applicants were sent to 658 firms.

Male applicants with AI capital received an interview invitation in 54% of cases, whereas male applicants without AI capital were invited to interview in 28% of cases.

Female applicants with AI capital received an interview invitation in 50% of cases, whereas female applicants without AI capital received one in 32% of cases.

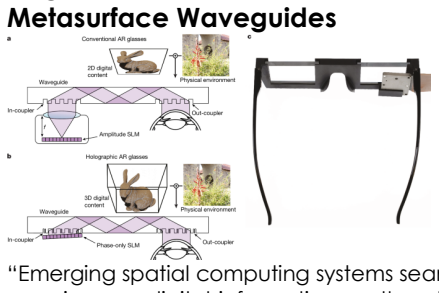
In large firms, applicants with AI capital were 36 percentage points more likely to be invited to an interview than in small-medium sized firms.

Male applicants with AI qualifications were shortlisted for jobs offering wages that were, on average, 12% higher than those for male applicants without AI capital, while female applicants with AI qualifications were offered interviews for jobs offering wages that were, on average, 13% higher than without AI capital."

Source: ARU (10 May 2024)

AR

## Full-Colour 3D Holographic Augmented-Reality Displays with Metasurface Waveguides




"Emerging spatial computing systems seamlessly superimpose digital information on the physical environment observed by a user, enabling transformative experiences across various domains, such as entertainment, education, communication and training. However, the widespread adoption of augmented-reality (AR) displays has been limited due to the bulky projection optics of their light engines and their inability to accurately portray three-dimensional (3D) depth cues for virtual content, among other factors4,5. Here we introduce a holographic AR system that overcomes these challenges using a unique combination of inverse-designed full-colour metasurface gratings, a compact dispersion-compensating waveguide geometry and artificial-intelligence-driven holography algorithms. These elements are co-designed to eliminate the need for bulky collimation optics between the spatial light modulator and the waveguide and to present vibrant, full-colour, 3D AR content in a compact device form factor. To deliver unprecedented visual quality with our prototype, we develop an innovative image formation model that combines a physically accurate waveguide model with learned components that are automatically calibrated using camera feedback. Our unique co-design of a nanophotonic metasurface waveguide and artificial-intelligence-driven holographic algorithms represents a significant advancement in creating visually compelling 3D AR experiences in a compact wearable device."

Source: Nature (8 May 2024)

ARCHITECTURE

## Architecture As a Product: What Makes a Building Worth Repeating?

"Architecture praxis has traditionally focused on customized, project-based services – a well-established model continually favored by industry professionals. While this approach yields remarkable built environments, it struggles to achieve scalability and longevity. The architectural industry has shown little interest in exploring alternative practices, processes and business models, considering that the traditional mode is equally vulnerable to market cycles as other industries are. Bespoke design solutions, the cornerstone of conventional practice, make standardizing processes and scaling services challenging. This focus also leads to fragmented workflows among firms and stakeholders. The stability and establishment of traditional practices can breed a risk-averse culture, hampering disruptive innovations within the industry.

The "productization" of architecture is an emerging concept that deviates from traditional forms of practice, viewing buildings as repeatable, scalable products. This approach augments the real estate value chain by fragmenting the design process. McKinsey & Co.'s "The Next Normal in Construction: How Disruption Is Reshaping the World's Largest Ecosystem" identifies product-based approaches as one of the likely forces shaping the industry's transformation in the coming decades. Many startups are experimenting with modular technology and component systems to productize buildings."

Source: Archdaily (6 May 2024)

ARCHITECTURE

## Roofs Of Trade: A Look At 12 Public Market Architectures

"Public Markets are complex cases that beautifully combine an engineering exploration with the instigation of a beneficial public function. In fact, it doesn't take much to designate an open and accessible commercial hub that would encourage commercial exchange and economic development for small or local businesses. All one needs is a lot and a sheltering roof. However, many designers and architects have chosen to take this exercise beyond the social level and into further material and technical exploration.

As in many cases, the public market's ground space is purposefully left unpartitioned and vast for flexible use and organic segmentation. The floor belongs to the merchants that occupy it differently. This requires the designers to address the structural challenge of ensuring wide areas between the columns. Many have taken this opportunity to develop vaulted wooden roofs, brick roofs that act as light tunnels, or even some that might suggest optimal circulation within a city or town.

For centuries, this public function has opened the door for some interesting and aesthetically unique design interventions, as we see in the global projects listed below."

Source: Archdaily (10 May 2024)

DESIGN

## Design No Longer "Just About Coming Up with Ideas and Making Stuff" says Oki Sato

ELECTRONIC WASTE

## These Recyclable Circuit Boards Could Stem E-Waste: Vitrimer-Based Pcbs Can Be Broken Down into A Gel for Repeated Reuse

MEDICAL TECH

## 'Smart' Contact Lenses Could Someday Enable Wireless Glaucoma Detection

ROBOTICS

## Robotic System Feeds People with Severe Mobility Limitations

"The pressures of sustainability are making it harder to be a young designer starting out in the industry, Nendo founder Oki Sato tells Dezeen in this interview.

Speaking to Dezeen at the Milan design week launch of his Hana-arashi collection for Italian brand Paola Lenti, Sato noted that it was just over 20 years since he first showed his work in Milan.

The designer, who founded architecture and design studio Nendo in 2002, showed as part of Salone Satellite – the showcase for young designers at Milan furniture fair Salone del Mobile – in 2003.

"It's very tough for a young designer."

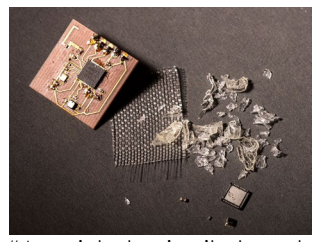
Sato believes he would find it harder to start out in the industry today, as there are more demands placed on designers to focus on sustainable design.

"Year after year I think the responsibility of the designers is getting very tense – it's huge, and I feel that it would be really difficult for me to start as a young designer in this period because design is not just about coming up with interesting ideas and making stuff," Sato told Dezeen.

"You have to think about materials and the process – is it not just human-centred, but for the planet, and we have to think about how it will be recycled in the future as well," he added.

"It's very tough for a young designer or a small atelier or workshop to think about all of these issues.""

Source: Dezeen (9 May 2024)

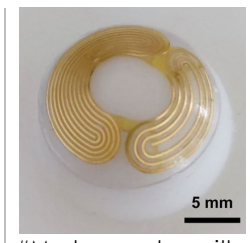


"A printed circuit board (PCB) made of materials that can be repeatedly recycled could help alleviate the world's growing electronic-waste problem. The recyclable PCBs, described in a paper published 26 April in the journal Nature Sustainability, perform as well as conventional boards, but unlike conventional PCBs they can be readily repaired. Additionally, the materials used to make recyclable PCBs can be recovered and reused.

According to the United Nations, the world produced 62 billion tonnes of electronic waste in 2022, and we're on track to produce 82 billion tonnes in 2030—an increase of over 30 percent. Chemicals can leach out of electronic waste and pollute water supplies; when e-waste is burned to recover valuable materials like gold and copper, it produces air pollution.

Conventional PCBs, which are the substrates that hold computer chips and their interconnections, are at the heart of this environmental problem. They are made up of glass fibers mixed with a type of plastic called a thermoset. These polymers look like a pile of cooked spaghetti. When they're heated in the process of strengthening the board, all the points where the polymers cross one another become chemically bonded. Those chemical bonds are not reversible, and so conventional PCBs are not recyclable. They also can't be readily repaired when they crack."

Source: IEEE Spectrum (2 May 2024)




"Most people with early-stage glaucoma don't know they have it, even though early treatment is key to reducing vision loss. While detecting a subtle increase in eye pressure helps doctors to diagnose glaucoma, it's challenging to monitor continuously, especially with the variety of temperatures eyes experience. Now, researchers in ACS Applied Materials & Interfaces report a prototype "smart" contact lens that measures eye pressure accurately, regardless of temperature…

First, Xiao and the team designed two miniature spiral circuits, each with a unique natural vibration pattern that would change when stretched by minute amounts, such as with changes to an eye's pressure and diameter. To create pressure-detecting contact lenses, the researchers sandwiched these tiny circuits between layers of polydimethylsiloxane, a typical contact lens material. Then they wirelessly read the embedded circuits' vibration patterns by holding a coil near the lens that was connected to a computer. The transmitted signals were unaffected by tests meant to mimic eye movement, extended exposure to moisture (to simulate damp conditions in the eye), and daily wear and tear."

Source: ACS (9 May 2024)

"Cornell researchers have developed a robotic feeding system that uses computer vision, machine learning and multimodal sensing to safely feed people with severe mobility limitations, including those with spinal cord injuries, cerebral palsy and multiple sclerosis.

"Feeding individuals with severe mobility limitations with a robot is difficult, as many cannot lean forward and require food to be placed directly inside their mouths," said Tapomayukh "Tapo" Bhattacharjee, assistant professor of computer science in the Cornell Ann S. Bowers College of Computing and Information Science and senior developer behind the system. "The challenge intensifies when feeding individuals with additional complex medical conditions."

A paper on the system, "Feel the Bite: Robot-Assisted Inside-Mouth Bite Transfer using Robust Mouth Perception and Physical Interaction-Aware Control," was presented at the Human Robot Interaction conference, held March 11-14, in Boulder, Colorado. It received a Best Paper Honorable Mention recognition, while a demo of the research team's broader robotic feeding system received a Best Demo Award."

Source: CORNELL (8 May 2024)

To view past Weekly Alerts CLICK HERE
For more articles or in-depth research, contact us at library@sutd.edu.sg