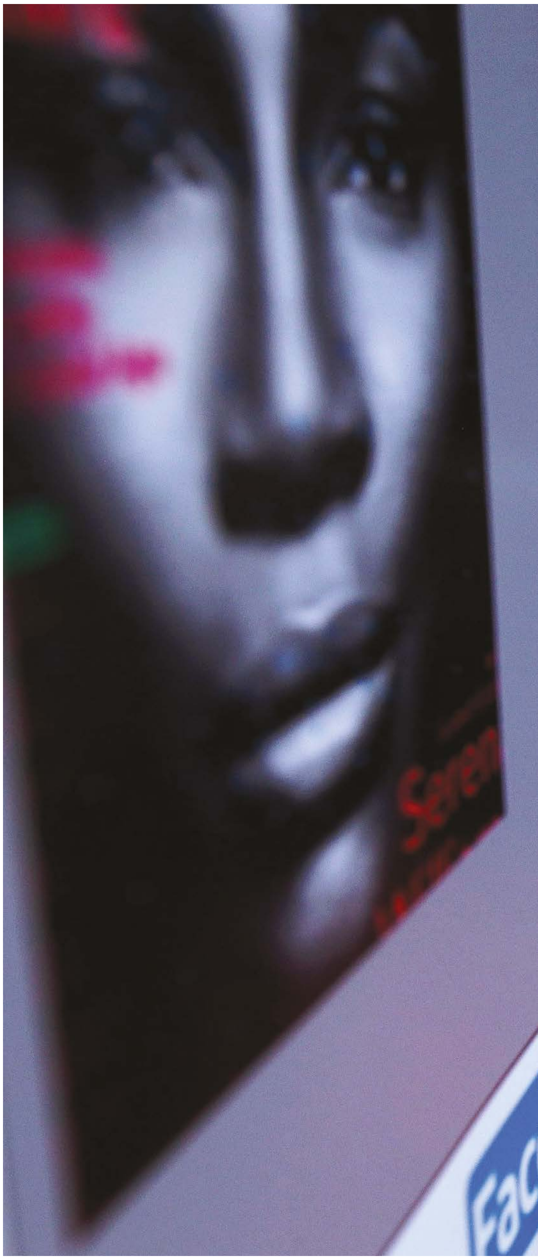
Reading between the lines

From facial recognition to drug discovery, these emerging technologies are the ones to watch.



New applications powered by artificial intelligence (AI) are being embraced by the public and private sectors. Their early uses hint at what's to come.

Recognizing bias
Facial recognition

In June 2020, IBM, Amazon and Microsoft announced that they were stepping back from facial-recognition software development amid concerns that it reinforces racial and gender bias. Amazon and Microsoft said they would stop selling facial-recognition software to police until new laws are passed in the United States to address potential human-rights abuses. IBM has called for an industry-wide review of the technology and how it is used.

Earlier in the year, Sundar Pichai, chief executive of Google's parent company, Alphabet, backed a proposal by the European Union to temporarily ban the technology in public places, such as train stations and stadiums.

> **"There is a realization that there are high-stakes consequences, especially for marginal communities."**

Facial-recognition systems are trained using a vast number of images to create 'faceprints' of people by mapping the geometry of certain facial features. Faceprints are used to classify a face into categories such as gender, age or race, and to compare it to other faceprints stored in databases. According to a 2018 report by the US National Institute of Standards and Technology (NIST), between 2014 and 2018, facial-recognition systems became 20 times better at finding a match in a database of 12 million portrait photos. But a separate study, published by NIST in December 2019, found that African-American and Asian faces were misidentified 10 to 100 times more often than Caucasian men (P. J. Grother *et al.* NIST Interagency/Internal Report 8280; 2019). The systems also had difficulty identifying women.

In a seminal paper published in 2018 by Timnit Gebru, former co-lead of Google's Ethical AI Team in Mountain View, California (see page S116) and Joy Buolamwini from the Massachusetts Institute of Technology (MIT) Media Lab in Cambridge, Massachusetts, male and light-skinned subjects were found to be more accurately classified than female and dark-skinned subjects. The study states that such errors occur largely because the systems were trained using mostly white, male-dominated data sets. One widely used data set was roughly 75% male and 80% white (J. Buolamwini and T. Gebru *Proc. Mach. Learn. Res.* **81**, 77–91; 2018).

"There is a realization in the computer facial recognition community that there are high-stakes consequences, especially for marginal communities," says Ellen Broad, an AI governance and ethics expert at the Australian National University in Canberra.

A Nature survey, featured in a November 2020 special issue on facial recognition, revealed discomfort among some researchers working in the field, particularly with regards to how its use might affect minority groups (see *Nature* **587**, 354–358; 2020). The 480 respondents said they were most uncomfortable about the use of facial recognition for live surveillance in schools and workplaces, and by companies to monitor public spaces, but were generally in support of it being used by police in criminal investigations.

The New York Police Department (NYPD) has been using facial-recognition technology since 2011 to match faces of unidentified criminals in surveillance footage and crime scene photos to those on watch lists. The NYPD states that no known case in New York City involves a person being falsely arrested on the basis of a facial-recognition match.

Despite concerns, facial-recognition software use is on the rise in many countries. It is now widespread in China and its use is growing rapidly in India. In 2019, the Australian government's Identity-matching Services Bill authorized the Department of Home Affairs to create and maintain facilities for the sharing of facial images and other identity information between government agencies and, in some circumstances, private organizations. **Leigh Dayton**

Blink and you'll miss it
Deepfake

The origins of deepfake, using AI to create falsified photos, videos or audio files, can be traced to 2017, when pornographic videos spliced with celebrities' faces were posted online. The early iterations that followed were relatively easy to

Joy Buolamwini from the MIT Media Lab says facial-recognition software has the highest error rates for darker-skinned females.

STEVEN SENNE/AP/SHUTTERSTOCK

detect, says Siwei Lyu, professor of computer science at the State University of New York at Buffalo, who pioneered a technique when he noticed that people in deepfake videos blinked far less often than they do in real life. As the technology advanced, so did Lyu's detection algorithms. They are now designed to identify the subtle changes left by creators when they resize or rotate a person's face to merge or superimpose it onto an image or video.

One way to create a deepfake video, Lyu explains, is using a neural network called a generative adversarial network (GAN), split into two parts. The 'generator' is trained using thousands of photos, videos or audio files to create hyper-realistic, but false, versions. The 'discriminator' detects poorly made fakes by comparing them to the real thing. The two parts might compete back and forth millions of times before the generator creates something realistic enough to 'fool' the discriminator.

The more advanced the GAN, the less training data it requires. That's one reason why Hao Li, professor of computer science at the University of Southern California in Los Angeles, thinks the battle to detect deepfakes could already be lost. "The fakes are getting so much better that they're undetectable in real time," he says.

In 2019, *The Wall Street Journal* reported that deepfake audio was used to impersonate the voice of a chief executive from an unnamed UK-based energy firm to complete a fraudulent €220,000 (US$261,000) bank transfer.

Li hopes that the more people learn about the existence of deepfakes, the less effective they will become. In September 2020, Microsoft Research launched its Video Authenticator software to combat the spread of disinformation through deepfakes, particularly those designed to undermine electoral processes and COVID-19 health advice.

Although there is an urgent need to combat deceptive deepfakes, the technology is increasingly being embraced by researchers, particularly in fields such as medicine and astronomy.

In 2019, for example, a team from the Institute of Medical Informatics at the University of Lübeck in Germany proposed the use of GANs to create hyper-realistic medical images to train AI used in cancer diagnosis. And a team led by the Lawrence Berkeley National Laboratory in California is using its 'CosmoGAN' system to produce maps of dark matter in the Universe. **Bill Condie**

Quicker treatments

AI drug discovery

In January 2020, researchers launched the world's first human clinical trial of a drug discovered using AI. Intended to treat obsessive-compulsive disorder (OCD), the compound DSP-1181 was identified using an AI drug discovery platform called Centaur Chemist.

After trawling through chemical libraries containing thousands of molecules, Centaur Chemist found the most relevant compounds for regulating serotonin, a chemical in the brain that has been linked to OCD. These were synthesized and tested in the lab before DSP-1181 was selected for a clinical trial in Japan.

The project is led by Exscientia, the British pharmaceutical company that owns Centaur Chemist, and Sumitomo Dainippon Pharma, a Japanese pharmaceutical firm. It took less than 12 months for the team to progress from the discovery stage to the end of preclinical testing. The global average is four to six years. If DSP-1181 obtains regulatory approval, it would be a major feat. Ninety per cent of compounds that start phase I trials (the earliest trials of drugs in people) fail to make it to market.

AI-powered systems have the potential to rapidly accelerate the drug-discovery process and thereby reduce spending. Bringing a new drug to market costs, on average, up to US$2.8 billion, and can take more than a decade. "Drug discovery is so expensive; anything, no matter how small, is going to be of immense benefit," says Toby Walsh, professor of artificial intelligence at the University of New South Wales in Australia. "I can't think of a more pressing time to develop drugs more cheaply and quickly."

The project follows work by Hong Kong-based biotechnology company, Insilico Medicine, and its machine-learning platform, GENTRL. A 2019 *Nature Biotechnology* paper led by Insilico founder Alexander Zhavoronkov describes how, in 46 days, the company designed, synthesized and experimentally validated a drug that targets a protein linked to a number of diseases, including fibrosis (A. Zhavoronkov *et al. Nature Biotechnol.* **37**, 1038–1040; 2019). The drug has been found to be effective in rats.

It's unsurprising that pharmaceutical giants are increasingly partnering with AI firms for drug development. In August 2020, Exscientia was named alongside Pfizer, Bayer and Merck as one of 37 partners in the new Corona Accelerated R&D Europe initiative, the largest undertaking of its kind to discover and develop COVID-19 treatments. **Leigh Dayton**

Emotional intelligence

Reading non-verbal cues

Emotion AI describes systems that can recognize, respond to and simulate moods and emotions. Designed to detect non-verbal cues, including body language, facial expressions and tone of voice, technologies based on emotion AI are being developed by industries such as recruitment, health care and education.

In offices or classrooms, for example, emotional cues could be used to detect those who are overworked or disengaged, says Simon Lucey, director of the Australian Institute of Machine Learning at the University of Adelaide, South Australia. "It's a fantastic application base to flag early signs of people who are stressed," says Lucey. Researchers are also exploring how emotion AI could detect signs of depression and anxiety.

Research on emotion AI has grown rapidly. According to the Dimensions database, the number of publications that mention 'emotion AI', 'sentiment analysis' and 'opinion mining' has increased almost tenfold between 2011 and 2019, from 1,225 publications to 11,926.

Companies such as Affectiva, founded by researchers from the MIT Media Lab, offer products aimed at detecting consumer reactions to brands and advertisements. In South Korea, nearly one-quarter of the top 131 corporations say they are using or plan to use emotion AI in recruitment, according to the Korea Economic Research Institute. Companies selling these systems, such as Midas IT, based in Pangyo, South Korea, claim that they can determine if a candidate is more likely to have qualities such as honesty or tenacity, based on expressions and word choice during interviews.

Researchers are calling for caution. They point to a lack of evidence that current emotion AI products can accurately deduce an individual's emotional state. Lisa Feldman Barrett, professor of psychology at Northeastern University in Boston, Massachusetts, says it is also concerning that such technologies assume that certain facial movements are universally perceived as expressions of particular emotions. Her research, which includes studies conducted with small communities in the Pacific, found diversity, rather than uniformity, in how people across cultures interpret facial movements. **Bill Condie**

MARK SOMMERFELD/THE NEW YORK TIMES/REDUX/EYEVINE

A lab technician at Deep Genomics in Canada, where AI is used in drug development.