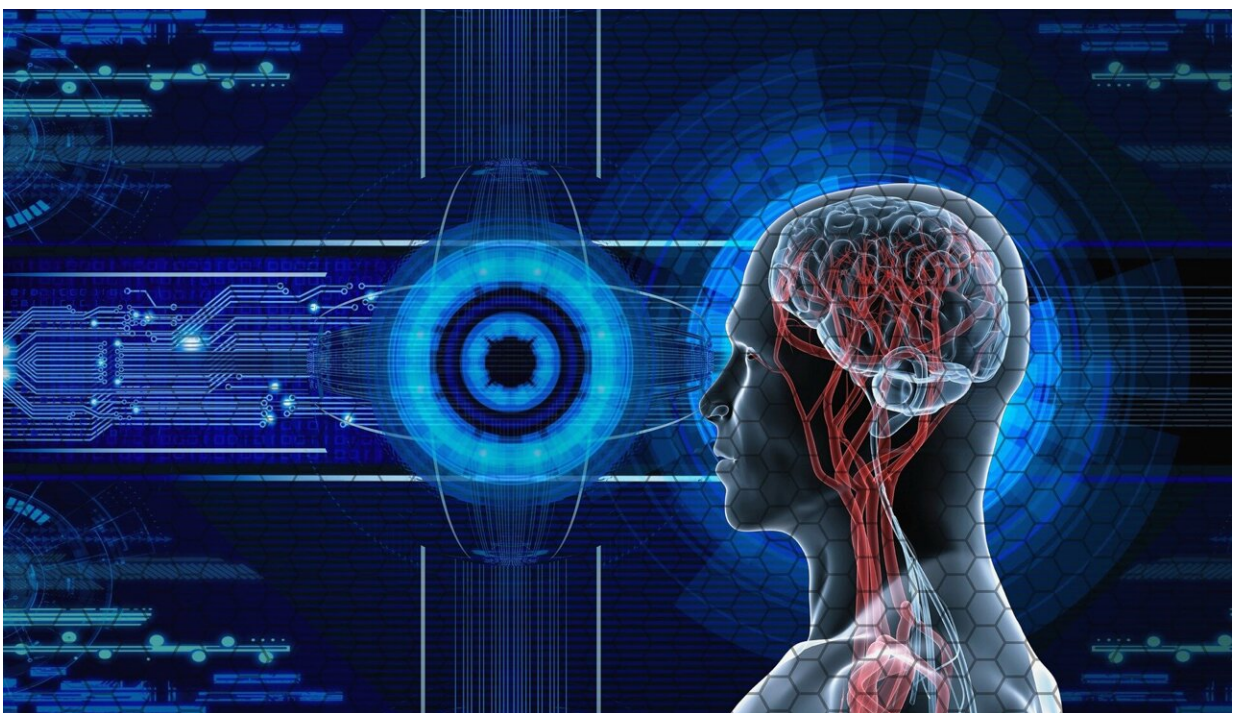


Researchers are tapping into psychology to improve AI to help robots interact more like humans

October 2 2023, by Geoff McMaster



Credit: Pixabay/CC0 Public Domain

A robot rushes down a busy hospital corridor dodging random foot traffic. With a subtle gesture from a care provider, the robot enters a room and hands over drugs to a waiting bedside nurse.

It may sound like a futuristic scenario, but an engineering professor at the University of Alberta, in collaboration with colleagues in psychology, is working to make it a reality within the next few years.

Ehsan Hashemi, with help from Dana Hayward and Kyle Mathewson in the Department of Psychology, is programming robots to work side by side with humans in dynamic work environments by responding to cues in body language.

Hashemi is already becoming well known for his work developing an artificial intelligence system for autonomous vehicles, but this time he's turning to expertise in a branch of experimental psychology called [human interactive cognition](#) to help robots interact more like humans.

"Humans are not always, or even often, rational beings ... and predicting their complex behavior continues to elude researchers," says Hayward.

A better understanding of our interactions—gaze and gestures, focused attention, decision-making, language and memory—can help AI researchers predict "what an individual will do or say next," she says.

Imagine 10 or 20 robots swerving around [human](#) workers in a warehouse, moving heavy materials at high speed. One hurdle of current navigation technology is that robots tend to stop in dynamic environments, "because they have no prediction over the [human movement](#) or that of other robots," says Hashemi.

"We're looking at how humans interact with each other with minimum information exchanged."

Hashemi and Mathewson have developed headsets with EEG sensors that, when worn by human workers, will feed brain-wave data into their predictive modeling, along with measurements of eye movement and

other body language. It's research that will move robots one step closer to interacting like human beings.

Intertwined from the start

Although the relationship between artificial intelligence and psychology seems like frontier science, it's one that has existed since the birth of AI. The neural networks first developed by originators of the technology in the 1950s were attempts to replicate the human brain.

Terms like "intelligence" and "deep learning" seem inseparable from our conception of human consciousness, with all of its strengths and potential failings. And as much as AI can help us understand more about human psychology and its disorders, psychology can also inform algorithms in ways that improve their functioning—while giving them the power to be dangerously manipulative.

Research exploring both sides of that equation—tapping our understanding of psychology to improve AI as well as interrogating its ethical, social and cultural implications—is expanding rapidly at the U of A.

In addition to AI's much-hyped potential for making our lives better, there is also a growing fear that it could exploit psychology in ways that flaunt our attempts at control. That anxiety is reflected in the recent [declaration by leading AI researchers](#) warning of a risk of extinction on par with nuclear war and global pandemics. The letter cites the threat of rampant disinformation, discrimination and impersonation.

Professor Geoffrey Rockwell, an expert in the burgeoning field of digital humanities, acknowledges AI's deep roots in psychology, prompting an [ongoing dialog](#) between our understanding of the human brain and the development of machine learning.

"Ideas about the brain influenced new designs for AI, and then those new designs influenced our understanding of the brain," he says.

Far beyond replicating or even exceeding the [human brain's](#) computational capacity, today's AI is taking on characteristics associated with human consciousness and behavior, if not actual sentience. In a [review published last year](#) in *Frontiers in Neuroscience*, the authors found that the predominant direction of AI research is to "give computers human advanced cognitive abilities, so that computers can recognize emotions, understand human feelings, and eventually achieve dialog and empathy with humans and other [artificial intelligence](#)."

In other words, the rational thinking of "brain" is now accompanied by the perceptual thinking of "heart."

An empathetic companion for the lonely?

One example is a project led by U of A computing scientist Osmar Zaiane. With a growing number of seniors suffering from loneliness, he is [exploring ways](#) with colleagues in psychiatry to create an empathetic and emotionally intelligent chatbot companion.

"An elderly person can say, 'I'm tired,' or, 'It's beautiful outside,' or tell a story about their day and receive a response that keeps them engaged," Zaiane says.

"Loneliness leads to boredom and depression, which causes an overall deterioration in health. But studies show that companionship—a cat, a dog or other people—helps tremendously."

But Zaiane also insists on carefully placed ethical guardrails. The chatbot can't offer most advice, beyond perhaps suggesting a sweater if the user mentions being cold, and it refrains from offering opinions, limiting

conversation to neutral topics such as nutrition, family and friends.

"The companion is relatively limited in what it can do," he says.

It's also designed to detect signs of depression and dementia, passing the information on to caregivers and health-care providers.

"If we detect anxiety and the possibility of self-harm, the bot might advise the person to call 811 or someone else for assistance." Anything beyond that, he argues, could be emotionally volatile and dangerous.

In the humanities, music professor Michael Frishkopf and his interdisciplinary research team are using machine learning to create music playlists and other soundscapes to [reduce stress in intensive care patients](#).

High stress levels, and anxiety associated with delirium and sleep deprivation, are common in critically ill patients, often compromising recovery and survival, says Frishkopf. Using drugs to treat these conditions can be expensive, often with limited effectiveness and potentially serious side-effects.

Frishkopf's "smart" sound system reads physiological feedback such as heart rate, breathing and sweat-gland response to customize calming sounds for individual patients. An algorithm essentially assesses a patient's psychological state, responding with a personalized playlist of soothing sounds.

The sonic prescription might also be matched to an individual's demographic profile, including gender, age and geographical background.

"Maybe the sounds you heard as a child or your musical experience

could have some special trigger for you," says Frishkopf.

A powerful diagnostic tool

Artificial intelligence is now also being used as a powerful tool for helping to diagnose mental disorders. By using AI to analyze brain scans, Sunil Kalmady Vasu, senior machine learning specialist in the Faculty of Medicine & Dentistry, and his research team have [found a way to assess the chances](#) that relatives of those with schizophrenia will develop the disease.

First-degree relatives of patients have up to a 19 percent risk of developing schizophrenia during their lifetime, compared with the general population's risk of less than one percent.

Though the tool is not meant to replace diagnosis by a psychiatrist, says Kalmady Vasu, it does provide support for early diagnosis by helping to identify symptom clusters.

To help doctors [diagnose depression](#), another U of A project goes beyond brain scans to include [social factors](#) in its data set.

"We don't have a clear picture of exactly where depression emerges, although researchers have made substantial progress in identifying its underpinnings," says project leader Bo Cao, an assistant professor in the U of A's Department of Psychiatry.

"We know there are genetic and brain components, but there could be other clinical, social and cognitive factors that can facilitate precision diagnosis."

Using data from the U.K. Biobank, a biomedical database containing genetic and health information for half a million people in the United

Kingdom, the researchers will be able to access health records, brain scans, social determinants and personal factors for more than 8,000 individuals diagnosed with major depressive disorder.

In computing science, researchers have successfully [trained a machine learning model](#) to identify people with post-traumatic stress disorder by analyzing their written texts—with 80 percent accuracy.

Through a process called sentiment analysis, the model is fed a large quantity of data, such as a series of tweets, and categorizes them according to whether they express positive or negative thoughts.

"Text data is so ubiquitous; it's so available and you have so much of it," says psychiatry Ph.D. candidate and project lead Jeff Sawalha. "With this much data, the model is able to learn some of the intricate patterns that help differentiate people with a particular mental illness."

Exploring the ethical implications

The U of A also has no shortage of scholars in the humanities and social sciences paying close attention to the ethical and social implications of AI fast becoming an integral part of our lives.

Vern Glaser in the Alberta School of Business points out in a [recent study](#) that when AI fails, it does so "quite spectacularly.... If you don't actively try to think through the value implications, it's going to end up creating bad outcomes."

He cites Microsoft's Tay as one example of bad outcomes. When the chatbot was introduced on Twitter in 2016, it was revoked within 24 hours after [trolls taught it to spew racist language](#).

Another example is the "robodebt" scandal of 2015, when the Australian

government used AI to identify overpayments of unemployment and disability benefits, in a sense removing any sense of empathy or human judgment from the equation. Its algorithm presumed every discrepancy reflected an overpayment and identified more than 734,000 overpayments worth two billion Australian dollars (C\$1.8 billion).

The human consequences were dire.

Parliamentary reviews found "a fundamental lack of procedural fairness" and called the program "incredibly disempowering to those people who had been affected, causing significant emotional trauma, stress and shame," including at least two suicides.

"The idea was that by eliminating human judgment, which is shaped by biases and personal values, the automated program would make better, fairer and more rational decisions at much lower cost," he says.

To prevent such destructive scenarios, human values need to be programmed in from the start, says Glaser. For AI designers, he recommends strategically inserting human interventions into algorithmic decision-making, and creating evaluative systems that account for multiple values.

"We want to make sure we understand what's going on, so the AI doesn't manage us," he says. "It's important to keep the dark side in mind. If we can do that, it can be a force for social good."

For Rockwell, a more immediate problem than the prospect of human extinction is the exploitation of human psychology to influence people in sinister ways, such as election interference or scamming seniors out of their savings.

He cites the [Cambridge Analytica scandal](#), in which a British political

consulting firm harvested the Facebook data of tens of millions of users to target those with psychological profiles most vulnerable to certain kinds of political propaganda.

The fear of such nefarious manipulation harks back to the alarm bell Marshall McLuhan sounded more than 50 years ago. McLuhan also warned that advertising could influence us in unconscious ways, says Rockwell.

"It turns out he was partly right, but advertising doesn't seem to work quite as well as people thought it would.

"I think we will also develop a certain level of immunity (to AI's manipulations), or we'll develop forms of digital literacy that prevent us from being scammed quite as easily as people worry we will be."

What we can't so easily resist, Rockwell argues, is the influence of human bias in AI's algorithms, given that it's a direct reflection of our historical, social and cultural conditioning.

"I don't think it's possible to eliminate bias from any data set, but we can be transparent about it," he says, by identifying, documenting and eliminating what we can.

"With data sets there has been this sort of land grab, where people just snarfed up data without asking permission, dealing with copyright or anything like that," he says.

Now that we know it's a problem, "we may see slower, more careful projects that try to improve the data."

Provided by University of Alberta

Citation: Researchers are tapping into psychology to improve AI to help robots interact more like humans (2023, October 2) retrieved 30 April 2024 from <https://techxplore.com/news/2023-10-psychology-ai-robots-interact-humans.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.