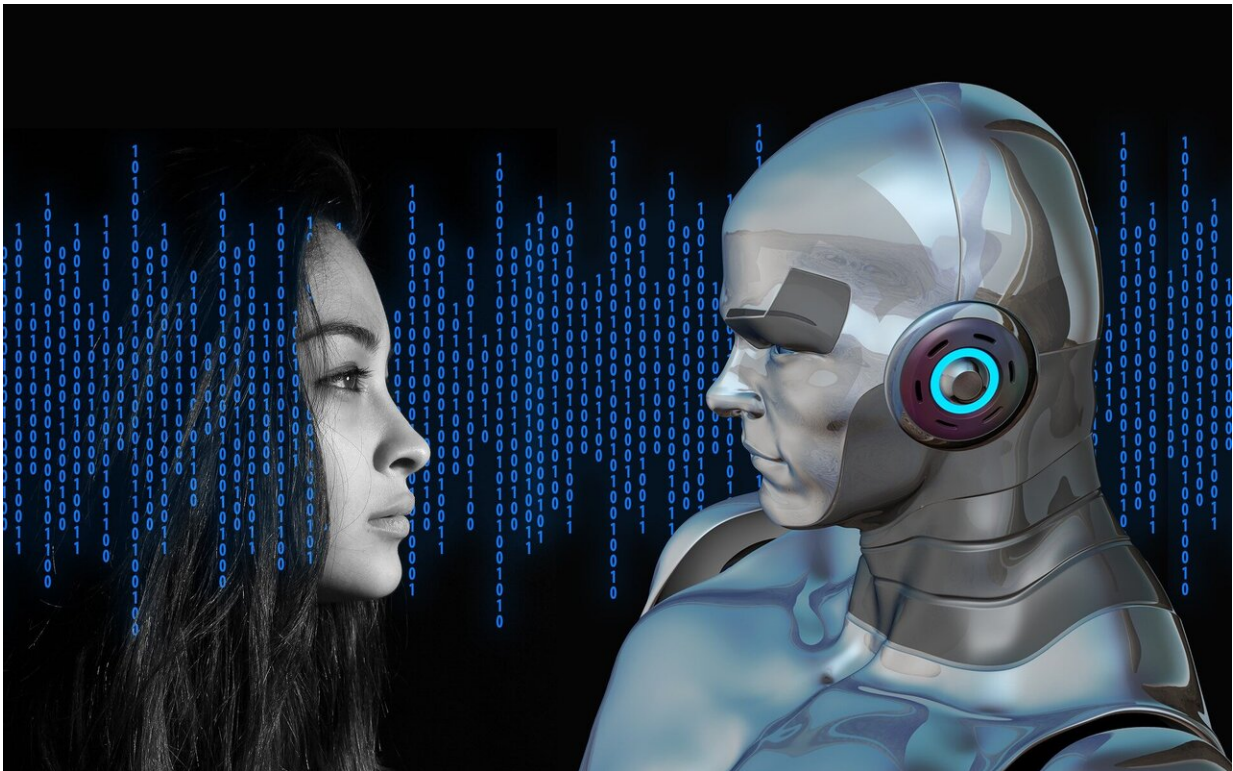# Machine sentience and you: What happens when machine learning goes too far

January 30 2024

Credit: CC0 Public Domain

There's always some truth in fiction, and now is about the time to get a step ahead of sci-fi dystopias and determine what the risk of machine sentience can be for humans.

Although people have long pondered the future of intelligent machinery, such questions have become all the more pressing with the rise of artificial intelligence (AI) and machine learning. These machines resemble [human interactions](): They can help problem solve, create content, and even carry on conversations. For fans of science fiction and dystopian novels, a looming issue could be on the horizon. What if these machines develop a sense of consciousness?

Researchers published their research into this topic in the [*Journal of Social Computing*]() on December 31, 2023.

While there is no quantifiable data presented in this discussion on artificial sentience (AS) in machines, there are many parallels drawn between human language development and the factors needed for machines to develop language in a meaningful way.

"Many of the people concerned with the possibility of machine sentience developing worry about the ethics of our use of these machines, or whether machines, being rational calculators, would attack humans to ensure their own survival," said John Levi Martin, author and researcher. "We here are worried about them catching a form of self-estrangement by transitioning to a specifically linguistic form of sentience."

The main characteristics making such a transition possible appear to be: unstructured deep learning, such as in [neural networks]() (computer analysis of data and training examples to provide better feedback), interaction between both humans and other machines, and a wide range of actions to continue self-driven learning. An example of this would be self-driving cars. Many forms of AI check these boxes already, leading to the concern of what the next step in their "evolution" might be.

This discussion states that it's not enough to be concerned with just the development of AS in machines, but raises the question of if we're fully

prepared for a type of consciousness to emerge in our machinery. Right now, with AI that can generate blog posts, diagnose an illness, create recipes, predict diseases or tell stories perfectly tailored to its inputs, it's not far off to imagine having what feels like a real connection with a machine that has learned of its state of being. However, researchers of this study warn, that is exactly the point at which we need to be wary of the outputs we receive.

"Becoming a linguistic being is more about orienting to the strategic control of information, and introduces a loss of wholeness and integrity…not something we want in devices we make responsible for our security," said Martin. As we've already put AI in charge of so much of our information, essentially relying on it to learn much in the way a [human brain](#) does, it has become a dangerous game to play when entrusting it with so much vital information in an almost reckless way.

Mimicking human responses and strategically controlling information are two very separate things. A "linguistic being" can have the capacity to be duplicitous and calculated in their responses. An important element of this is, at what point do we find out we're being played by the machine?

What's to come is in the hands of computer scientists as they develop strategies or protocols to test machines for linguistic sentience. The ethics behind using machines that have developed a linguistic form of sentience or sense of "self" are yet to be fully established, but one can imagine it would become a social hot topic. The relationship between a self-realized person and a sentient machine is sure to be complex, and the uncharted waters of this type of kinship would surely bring about many concepts regarding ethics, morality and the continued use of this "self-aware" technology.

More information: Maurice Bokanga et al, Through a Scanner

Darkly: Machine Sentience and the Language Virus, *Journal of Social Computing* (2024).