



The Artificial Intelligence Zombie



Untangling Anthropomorphism in Responsible Artificial Intelligence Development

Julien da Silva, BA Hons. Philosophy | P.I. Jocelyn Maclure

undergraduate poster showcase

INTRODUCTION



Fact #1:

We can expect to see even more advanced versions of **social robots** and artificial intelligence entering our everyday spaces as they continue to gain acceptance and popularity.

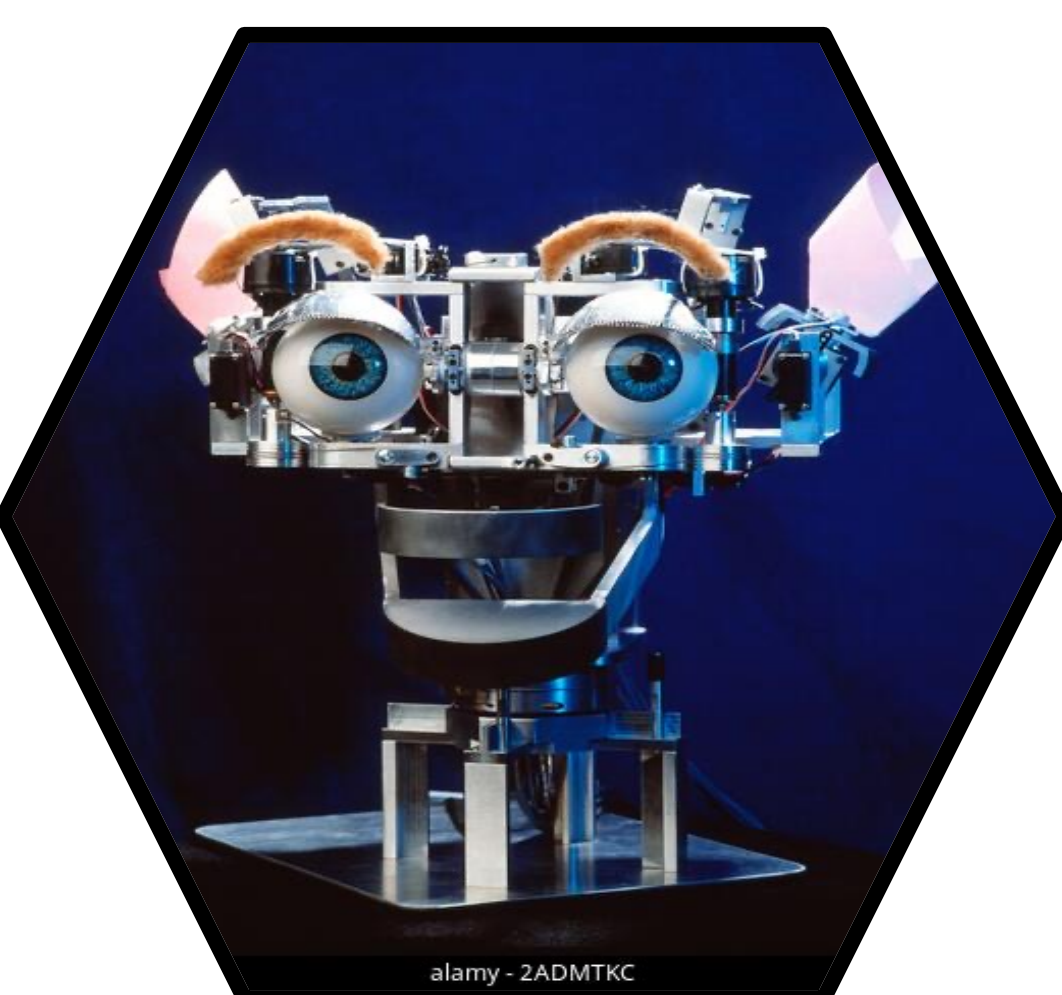
Pictured: Innvo Labs' PLEO robot dinosaur

Fact #2:

A **moral agent** is someone or something that can understand what is right and wrong and has the ability to make choices based on that understanding.



Pictured: Softbank Robotics' Pepper



Fact #3:

As humans, we tend to give **human-like qualities** and personalities to the objects we interact with.

Pictured: MIT Media Lab's Kismet

PROBLEM: As technology advances, people might start seeing human-like robots as conscious and capable of feeling emotions, **even when they're not**. This could lead to confusion and irresolvable debates around their treatment and legal status, potentially harming the development of artificial intelligence and the way we interact with it.

CAN HUMAN-LIKE ROBOTS TELL RIGHT FROM WRONG?

If so, how should we treat them?

Think about the ways we assign blame. Mature human beings understand right from wrong, and so we hold them accountable for their actions. On the other hand, babies and animals cannot, so it would be wrong to blame them morally accountable for their behaviour. Even so, they deserve our kindness, attention, and respect.

WHY DOES THIS MATTER?

1) It may look like some robots deserve rights, even if they don't.

If we imagine a future where human-like robots are common, some people might argue that they should have rights like humans. However, this would be a waste of political time and resources, as these robots don't actually have feelings or consciousness like we do. It's more important to focus **on protecting the rights of actual humans** before worrying about robots.

2) There are more urgent legislative priorities.

Both citizens and lawmakers alike need direction in order to differentiate between urgent and unimportant concerns. The advent of human-like artificial intelligence forces lawmakers to think about the speculative possibility of conscious artificial intelligence at a time when, in the short term, artificial intelligence threatens democracy, equality, and justice, just to name a few.

3) We need to know where to assign blame.

When human-like robots do something wrong, it can be hard to decide who's at fault, especially when it comes to legal matters. Sometimes, governments and companies might choose to blame the robot because it's easier than blaming a human. However, it's important to assign blame correctly in order to make sure that justice is done.

HOW SHOULD WE ADDRESS THE PROBLEM?

I suggest three potential avenues for addressing the problem:

1) Warning-focused approach

Warn consumers in advance that social robots and AI are unconscious.

2) Reverse-Turing test regulation

Impose a requirement that any commercial robot must clearly communicate that it is unconscious. In other words, a reasonable person must not be able to mistake a robot or artificial intelligence for a conscious being.

3) Non-proliferation or moratorium agreement

Greatly limit the number of commercially available human-like artificial intelligence systems, or put a moratorium on their commercial availability, thus avoiding a widespread adoption scenario.

Selected Bibliography

- Braun, Joachim von, Margaret S. Archer, Gregory M. Reichberg, and Marcelo Sánchez Sorondo, eds. 2021. *Robotics, AI, and Humanity: Science, Ethics, and Policy*. Cham: Springer International Publishing.
- Darling, Kate. 2016. "Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior towards Robotic Objects." In *Robot Law*, by Ryan Calo, A. Froomkin, and Ian Kerr, 213–32. Edward Elgar Publishing.
- Kahn, Peter, Takayuki Kanda, Hiroshi Ishiguro, Brian Gill, Jolina Ruckert, Solace Shen, Heather Gary, Aimee Reichert, Nathan Freier, and Rachel Severson. 2012. "Do People Hold a Humanoid Robot Morally Accountable for the Harm It Causes?"

Acknowledgements

Many thanks to The McGill Office of Science Education for the opportunity to present and the assistance provided in the process.

Contact: julien.dasilva@mail.mcgill.ca