

Can we be friends with AI?

What risks would arise from the proliferation of such friendships?



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Nick Munn (nick.munn@waikato.ac.nz) & Dan Weijers (dan.weijers@waikato.ac.nz)

**KO TE TANGATA
FOR THE PEOPLE**

The pitch: There is no good in-principle reason to oppose the development of human-AI friendships.

The structure:

- What matters for friendship?
- Do AI have that? (Will they eventually have it?)
- Is there something else that makes human-AI friendships bad/impossible? (no!)
- What follows from the above?

The important aspects of the concept of friendship



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

1. A preponderance of rewarding interactions
2. Mutual positive intentions

Friendship is a relationship constituted by a clear majority of rewarding interactions between parties with positive intentions toward each other.

You need both (1) & (2). If you have both, you have a friendship. The more of them you have, the better the friendship is.

Of course, Aristotelians would disagree...

Can AI be friends?



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

1. Do the AI we currently have, have the capacity for friendship?
2. What about potential near-future AI?
3. Or hypothetical future AI?

Maybe? (1). Yes (2). Yes (3).



We already have mutual positive intention:

- Replika users call their AI companions friends.
- AI can be programmed to make user wellbeing an end-goal.

...and rewarding interactions:

- People have been using (paying for!) these systems for years.
- AI can be programmed to recognise reward via praise or other positive interactions.



OBJECTION(s)!



The friendship won't feel real/valuable



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

But... it does, already. Petina, Hancock & Xie, 2023 found that deeply involved Replika users rated their AI companions as human, or better than human!

(some) Replika users (report that they) fall in love with their AI.

The more people interact with AI, the more they treat AI as human.



A fundamental Inequality?



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Are AI the right kinds of agents to be friends with us? The people making them want us to think they are!

Inequality between friends doesn't defeat friendship.

- Those with severe intellectual disabilities can be friends with us – these friendships are real, valuable and mutually beneficial.
- Other human-nonhuman friendships are real (even on an Aristotelian account if we have reasonable expectations about nonhuman reciprocity; Olbrey 2016)



AI lacks [something] relevant to friendship



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

But what? Humanity? Emotions? Empathy? Values?

Without this, is our so-called friendship merely a transactional façade?

Functionally, AI **do** have the relevant aspects. They don't need emotion, or sentiment.

Many human instances of caring for friends involve sentiment, but not all do. Sentiment predicts action, but so does intention.

Replika

The AI companion who cares

Always here to listen and talk. Always on your side.
Join millions growing with their AI friends now!

What would AI friendships displace?



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Real friendships! With real people!

Only if you have those... If not, AI friendships are clearly a net positive.

And... maybe AI friendships are (or can be) better, anyway?





What are the positives?

The advantages of AI friendship



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

AI can teach us how to be better friends.

AI can help people recover from traumatic experiences with humans (Maeng & Lee, 2021).

AI can overcome contingent disadvantages (bad luck) in the availability of human friends.

If we accept that human-AI friendships are real... why do we need humans any more?

- People want things from friendship: support, understanding of their circumstances, availability on the part of their friends...
- For at least some people, some of the time, AI friends are more capable of providing this than are human friends.
- Soon™ AI will be more capable of providing such support and understanding, for more people.

Of course, many people will continue to have highly rewarding friendships with other humans, and nothing we have said here provides any reason to abandon those friendships.

So...



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

There is no moral responsibility to discourage friendships between humans and AI.

They aren't obviously bad. They might (already) be good. In the future, they are even more likely to be positive.

Perhaps we need systems in place to protect friendships of this kind.

But... The only people who need to worry about AI friends are bad human friends!

Thank you



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Nick Munn

nick.munn@waikato.ac.nz

Dan Weijers

dan.weijers@waikato.ac.nz

