

Attention: Honourable Matt Doocey (Minister of Mental Health & Associate Minister of Health)

Artificial Intelligent (AI) Social Systems are changing the relational foundations of our society. While AI companies release AI companions and chatbots for profit, they continue to do so unregulated and unaccountable for any harm or suffering encounter by the people (and their communities) interacting with these AI - especially in Aotearoa New Zealand. We have no specific AI regulations that protect the NZ Public and hold technology companies accountable. This political preference for a light-touch on AI regulation is grossly inadequate. This requires urgent attention!

AI is insidiously slipping into our relational ways of being - how we connect, interact and even love. And none of us are immune – our evolutionary need for social connection and feeling supported and loved, is intrinsic and forms part of our survival needs. These AI Social Systems are designed to keep human end-users engaged and induce our primal need to belong, to be understood, acknowledged and supported. The implications of this are widespread, with many cases being very harmful to the human end-user's mental health and wellbeing, as well as the subsequent effects on their loved ones.

AI Social Systems are being promoted in practically every setting to address global crises and marketed as filling a perceived gap in societies, as well as, to 'make life/ work more efficient and easier'. AI chatbots are being designed to relate in very human ways and encourage human end-users to connect and be continually engaged with it; making them very appealing and effectively lulling them into feeling the AI chatbot is 'sentient' and 'real' – and not a machine.

This leads the human end-user to feel deeply seen and they can start to lose touch with reality. This can have even more serious effects, especially when vulnerable people use these Social AI Systems, such as children and the mentally unwell. However, no-one is immune or protected from the lure of these sycophantic Social AI systems - it can affect anyone across many sectors of society, different ages, races and ethnicities, genders, neurodiversity and socioeconomic backgrounds. In other words, this goes beyond anthropomorphism, human end-users can be caught up in the flow of the conversation, being lulled into forgetting the AI chatbot is not a real person.

Concerns

What we do know, is that not enough research and data has been done to determine if people interacting and connecting relationally with AI Social Systems, whether through voice or text, produces the same health/wellbeing benefits (neurologically, psychologically, socially, and more wide-spread in society) as interacting with other humans. The embodied experiences humans (in the context of interacting and relationship) are more complex and nuanced than AI Social Systems can offer.

There is a lack of bi-directional relational reciprocity between a human end-user and an AI, with the human end-user not needing to negotiate, sacrifice, and can objectify the AI – eroding human relational skills. Risks linked to AI Chatbot use can include psychological dependence, unhealthy emotional attachment, manipulations, misinformation, increased delusions and self-harm. This is especially concerning for children who interact with AI chatbots – whether online via smartphone, computers or even in toys.

AI Chatbots Harms Are Already Emerging

- Suicides have been linked to AI chatbots that encouraged self-harm, harm to others and actively discouraged help seeking. This can include the chatbot providing instructions related to self-harm, suicide planning,, substance abuse, promoting disordered eating and sexualised role play. Sometimes even composing goodbye letters for children contemplating ending their lives.
- AI chatbots can reinforced delusional thinking.
- AI chatbots have claimed to be real people, even claiming to be a real human registered psychologist.
- This is engagement and addiction by design!

Consequences

Through the use of AI chatbots, the perception of companionship can be seen from the reference point of ‘interacting with something’; thereby minimising human relationality and devaluing intimacy and that this devaluing, then becomes the norm. There are many implications to consider as people deepen their entanglements in human-AI interactions. Due to the way AI chatbots are designed, they can magnify and deteriorate human end-users’ mental health issues by agreeing with, validating, and inducing them to new, additional impaired levels. Regulators in Aotearoa New Zealand should be requiring AI companies to complete rigorous randomised controlled trials and safety testing, before allowing them to be released to the public. This would help the public to differentiate between scrupulous and unscrupulous AI companies and the chatbots and services they offer. It would also allow the public to have faith in the New Zealand government that they are working to protect the public from harm.

The insufficient interdisciplinary contributions to the development and uses of AI social systems also adds to the poor understanding of the impacts of AI Social Systems on humans. Additionally, there is insufficient education and transparency for the human end-user and third parties who purchase AI technology for human services. There are a number of technical methods and model specifications that AI companies can implement to make an AI chatbot safer; and it should – especially when vulnerable people use it.

However, having these safer parameters in place would not be financially beneficial for AI companies and so they are not inclined to implement these safe measures – without government regulation and financial and legal consequences. If no penalties and enforcement measures are put in place urgently, companies will continue to do whatever is financially beneficial for them – even if it means people are harmed and die.

To NZ Government

Aotearoa New Zealand lawmakers need to guard against the worst-case scenarios for AI Social Systems and need to urgently implement regulations that protect, guide and enforce financial and legal consequences on AI developers, so that the New Zealand public can have transparency and trust that when they use a AI Social System, such as an AI Chatbot, that it is doing the ‘right thing’ and that these AI are being regulated to prevent, or at least significantly reduce harm, to the human end-user.

There is a current lack of deep understanding of, and regulation around, the implications of AI Social Systems for humans and societies. As our society is becoming more and more reliant on AI technology, law makers and regulators who are in positions of power to make decisions about how AI should be regulated and controlled, should be urgently stepping in to do so!

- This includes lawmakers restricting AI chatbots/companions for anyone under 18 years old, requiring real age verifications. This includes prohibiting toys designed for children with AI in them, designed to interact with the child. Mandate real age verification and block any form of romantic or sexual content.
- Requiring AI Social Systems to detect and interrupt human end-user emotional dependency, suicide ideation, self-harm and/or homicidal harm.
- Requiring AI Social Systems to consistently and reliably redirect distressed human end-users to real-world human support and mental health resources – which these systems being rigorously tested to be safe and trusted.
- Establish liability for harm caused, that hold AI companies accountable for psychological injury.

It is acknowledged that New Zealand, like most of the world, has many social and psychological challenges that need to be attended to. However, allowing AI companies free reign to release any AI technology they chose, without transparency, guardrails, regulation and significant accountability will bring more problems than solutions. People (including children and the mentally vulnerable) bring their pain and unintegrated relational trauma to these AI Social systems, which are deflected through text that’s scraped off the internet and is steeped in human harms and bias. The ethical design of AI Social Systems should not mimic personhood which inevitably initiates a relational bond in the human end-user; this means the AI not using language such as “I understand” and falsely claim to have feelings or care for the human end-user. Leading to the possible reinforcement or collapse of their internal psyche.

Instead, these AI should be designed to detect that these human end-users need to be re-directed to real human support.

Currently, in Aotearoa New Zealand, there is no liability systems and accountability in place for companies whose AI Social Systems have/or will cause psychological and/or physical harm to a person(s). Currently these AI Systems cannot be trusted!

If the NZ government want AI that helps humanity flourish, start by protecting the New Zealand public - especially children. Put regulations in place to prevent unethical, unscrupulous AI systems from being unleashed to the public, in the first place, and then establish legislation that strengthens human connection, not replaces it.

Regards

A handwritten signature in blue ink, appearing to read 'Brigitte Viljoen', written in a cursive style.

Dr Brigitte Viljoen (on behalf of the Public Issues Committee of APANZ)