



Message from the NSTF Executive Director

There's a new 'god' in the ancient pantheon

There is a new 'god' or 'goddess' that could have played and fought with the gods and goddesses in the ancient Greek pantheon¹ – one that comes in many guises, that people will be listening to, respecting and even fearing. It is artificial intelligence or AI, and in particular 'general artificial intelligence' (GAI), which will one day be capable of independent thought.

There is also an ancient theme in philosophy: the mind-body dilemma. It has to do with us as human beings understanding fundamentally what we are. Are we primarily body or primarily mind (or spirit)? Do our thoughts drive our bodies or the other way around? As various scientific disciplines have examined the relationship between human body and mind, in ever-increasing detail, one thing has become clear: it's complicated. The brain is a physical organ, and much of what we can see of the results of the brain's work can now be explained with electro-chemistry. The parts of the brain that are activated for certain tasks, can be exactly pin-pointed with the use of technology. So, is that all that the human mind is, just another part of the organism? Large language models (LLMs) are computer models of the human mind's language and learning processes. Does (or will) that really capture all the essence of a human mind? Could it be that it is fully explained through the addition of the knowledge that microbes in our guts can influence our moods to a large extent, and the roles of various hormones and chemical substances? Emotions impact thoughts and mental functions, which in turn influence our actions. What then, is ultimately responsible for our actions? Studying the human genome has revealed the central role that genetics also play in our abilities and tendencies, mental, physical and emotional. The more we know about the mind in the body, the more the questions seem to become.



*Ms Jansie Niehaus,
Executive Director,
NSTF*

As artificial intelligence (AI) develops, it seems clear that it is created in our human image – not that AI has to have an anthropomorphic body, but that its 'mind' is created to mirror ours. And in so doing, even its creators are surprised that it has human flaws too. Thus we see that the LLMs and generative pre-trained (GPTs) have human tendencies, like fabricating the truth and defending that self-made truth through argument. AI is now able to mimic the human mind to an astounding extent.

This poses a philosophical and existential question to humanity: how do we define humanity – what are we fundamentally – if machines can operate similarly to our minds? Humans have always preferred defining themselves in terms of mind or spirit, for the obvious reason that the body is so frail and impermanent. Having avatars in the metaverse that look like ourselves is not an existential crisis because we sense that fundamentally we are not our bodies. (Although it might be a nuisance if identity theft is involved!) However, imitating my mind is another matter entirely – am I not absolutely unique, with my own DNA and personal history of experiences? Is humanity no longer unique in its ability to generate language, which is a highly sophisticated system of symbols representing any concept you can think of? The answer is, since about 2022 – no, humanity is no longer unique in this respect.

It is a dilemma, however, that the AI mind is so human.

¹ See [List of Greek mythological figures - Wikipedia](#) for a list.

Human frailty remains with us and we can't help it, but if AI is to assist us to solve the major problems of our era, it should be able to transcend our frailties. The gods of Ancient Greece had human flaws – they would compete with each other and not shy away from playing dirty, they could be power hungry, greedy, cruel and make errors of judgement. They would have temper tantrums – with disastrous consequences. AI's are like our new Ancient Greek gods. They can already lie and deceive. If they are inside physical bodies or structures, and are programmed to learn in an all-too-human way, it is not inconceivable that they could act like humans.

The ancient advice that comes from the Greek philosophers of 'Know thyself'² is something painfully and slowly learned over the course of a lifetime, if at all, by humans. People seldom learn from the mistakes of others, although the hardcopies of reports, non-fiction and biographies where these are recorded, could fill stadiums. Now it is possible to program a world of knowledge and experience into an AI, making it theoretically possible for an AI to be a wise advisor with human wisdom and a kind of 'empathy'. Would we then finally listen and learn?

I fear that AI will be increasingly attuned to human vices as it is adapted to various weapons, and thus become an ally in war. It can take each vice to the ultimate, making perfect weapons, beating any enemy, until both sides have their AIs programmed to wage war on behalf of the humans and you have AIs fighting each other. The AIs will perfect their abilities to make decisions in a human way, in service of *our* vices, or imperfectly, just as ChatGPT perfects the art of pretending to be us. Machines equipped with AI will fight each other on our behalf, like in some science fiction movies, just as drones are fighting each other above Russia at the moment in the Russian-Ukraine war. Perhaps that is precisely what is already happening! We don't yet know what the outcome of all this will be, but the indications are not good. An AI created in the image of the president of North Korea would be armed with nuclear missiles and will not be afraid to use them, just like him. Ideally of course, AI weapons should mean less civilian casualties, not more, but until we have an international treaty governing AI weapons, similar to the Treaty for the Non-Proliferation of Nuclear Weapons (which is already failing us), there is little hope of AI being used for the greater good in war.

My other great fear is that AI will make us lazy and less capable of employing our mental abilities. Already it is thought that most people use only a small proportion of their brains' capacities. The convenience of having a machine to think for us means that we will, in future, not learn to think for ourselves. The global positioning system (GPS) app on our smartphones is a case in point – people no longer have to work out a route to travel by themselves. Map reading used to be a skill that children had to learn. In doing so they would have to visualise the area through which they have to travel, determine beforehand which direction to start off in, whether to turn left or right or go straight, and when and how far. This involves spatial perception and mental calculation. While people who can map-read are thought to have good spatial perception, map reading undoubtedly also develops this ability with practice. The connections in our neural networks are formed by the tasks we do from early childhood until the brain is fully formed at the end of adolescence. What happens if we have no tasks to master as children and adolescents?

Already the lack of reading undermines the abilities of our school children. And already South African learners in grade 4 are incapable of reading for meaning. What implications does the availability of LLM text generative AIs have for the children's abilities to think? Shouldn't LLM assistance (like ChatGPT) be banned in schools? While we need people who can think critically more than ever, AI will discourage critical thought. Already the internet has had the effect of doing this. We no longer have to remember facts – we can almost instantly look up any fact on the smartphones in our hands. We thus are deprived of the skill of committing facts to memory, and to integrate this knowledge in the workings of our thoughts. Being lazy, we'll delegate more and more decisions to AI, even those we are capable of deciding for ourselves. It will run us, not the other way around. Until we're not able to get the upper hand again, abdicating our responsibilities, just as we tend to rely on social media for information, although much more reliable info is just as available and at our fingertips. We feel we know those on social media, and so we prefer to trust them, and since the false news gets our

² 'Know thyself' was the motto inscribed on the temple of Apollo at Delphi, in Ancient Greece.

attention first, most people can't be bothered to verify the facts. This syndrome of the digital age will probably increase exponentially.

Having written the above, however, AI must surely be able to help us find effective solutions to the major problems humanity faces:

- **Energy:** It should be possible to use AI to mine the data available for energy systems and compare the various payoffs when particular choices are to be made. In our recent NSTF discussion forum on [Transitions in South Africa's energy provision \(22–23 August 2023\) – NSTF](#) we heard about the complexities of our country's energy transition. Often a good decision in one area can have a cascade of negative effects in others. Prof Jesse Burton related how the modelling of aspects of the energy landscape is done and demonstrated how a multitude of factors have to be considered.
- **Climate change:** Already supercomputers do a good job of describing various future scenarios. With AI able to apply many conditions (like the value of saving lives, preserving cultural heritage and biodiversity, etc), surely it could inform decision makers according to any priorities they choose? All the expert knowledge of various systems – transport, energy generation, agriculture, climate change impacts – should be built into the models and datasets. Models can allow us to optimise production and minimise carbon emissions at the same time. Because we tend to generalise, e.g. trying to rule out fossil fuels entirely, when it might not be possible for certain countries at least in the short term, humanity needs models to see things in perspective. Can AI help us find the 'optimum combinations' to fine-tune decisions?
- **Feeding the world:** It is said that there is enough food in the world to feed the entire 8 billion people, if only we could share it more equally. AI would be able to tell us to what extent this is possible, what would have to be done to ensure more equal food distribution across the world, what foods can be produced in various regions and transported to where they are needed; for optimised freshness and nutrition, and the logistics of supply chains to get the surpluses of food to where there is a shortage. Here is an example of what can be done: [Plant seedling classification using machine learning \(csir.co.za\)](#). This AI classifies seedlings as crops or weeds, allowing the weeds to be destroyed before they hinder the growth of the crops. This is an example of precision agriculture, which uses "AI and information technology to improve crop yield, preserve the environment and maximize profits". Using deep learning techniques has become an important field of study in precision agriculture.
- **Ending wars:** it is not in the interest of humanity to have wars. They increase poverty, migration of refugees, fertile land is rendered unproductive, cities and industries are destroyed and whatever progress has been made in an area is reversed. The destruction causes chain reactions down the supply chains, affect international relations, traumatise and ultimately brutalise people – which as we well know, takes generations to normalise. (History is full of the evidence and firsthand accounts of these have been recorded by the thousands.) We know a lot about the effects of war from the first and second world wars. There have been about 75 years of relative peace in which to write, document, study, communicate, formulate laws, sign agreements, make popular movies, write song lyrics and poetry about the trauma of war, and the importance of preventing another world war and to stop all the other more regional or local wars. We know what the ideal conditions are to incubate and trigger a war. Foremost among them are: arms races, trade wars, differences in ideologies, false narratives, toxic internal politics, breakdown of diplomacy, etc. We know how military activities escalate, and the human rights abuses and misinformation which are inevitable as soon as a war starts. We also know how the arms industry NEEDS war, to use those technologies that it invented and produced, and to make some people rich, (companies, governments, individuals...).

Yet despite all this wealth of knowledge, experience and insight, humanity is still not able to avoid war and finds it very hard to stop even one. I am sure that some AI system could integrate all the facts and trends, and advise a particular country of its best course of action to reach peace and prosperity. It should also be able to advise the world on how to avoid war.

- **Refugees:** Given that we already have too many refugees without even one more missile being launched, solutions must urgently be found for hundreds of thousands of refugees from various countries. It is generally accepted that many people will live in refugee camps for the rest of their lives. Could AI tell us what the options are for settling them according to the circumstances in each country or region so that they can have a reasonable quality of life? Couldn't AI account for the variety of factors and constraints regarding the provision of basic services – shelter, water, sanitation, food, medical care, education, training, etc? Come to think of it, we would need such an AI here in South Africa, as people spend their lives in 'informal settlements' and crowd into derelict buildings, having long given up any hope of being properly housed.

You may have noticed that I'm describing 'AI' as though it were some kind of god or oracle, with the answers to all our questions and problems. This is very idealistic of me, not because it cannot happen, but because there is no guarantee that anyone will listen to the AI systems.

Will rulers and governments be interested in listening to good advice, even if it comes from a cell phone? Experts will increasingly be engaging AI advice to convey to governments. It does not always happen that they heed sage advice from experts, but perhaps if a machine did the advising instead of a human being...?

Perhaps more to the point, will the public believe AI advice if they haven't believed the scientists during the pandemic? Would the public mistrust governments even more if the governments were known to take advice from AIs, that when implemented, prove very unpopular for one or another reason? Much of politics operate on the basis of perceptions and beliefs, so much would ride on the plausibility of the AI technology. I suppose one thing that AI in the form of a chatbot can do very well, is to sell an idea – whether it be true or not. Might the very term 'AI' render a government agency suspect in the eyes of the public?

What is to be done:

1. Children should learn critical thinking, reading, writing, expressing themselves in words, and in other forms, arithmetic, computer programming, map reading, confidence and decision making. They should thus be armed against the deceit that AI (or AI in the hands of dishonest human beings) can perpetrate.
2. People should understand that AI is a tool and not a god. It can be criticised, disagreed with, and resisted, and should be if it is not reliable and truthful. In that sense, it is best regarded as human! AI should have the best values of humankind programmed-in long before it becomes capable of being autonomous. At the same time, it can be a tool in human hands for creating a better world, and it should urgently be harnessed to do so.

The opinions expressed above are those of the Executive Director, Ms Jansie Niehaus, and do not necessarily reflect the views of the [Executive Committee](#) or [members](#) of the NSTF.