Applying the Theological Brakes to AI and Techno-enthusiasm

Chris Mulherin

Abstract:

Like earlier products of science and technology, rapid advances in artificial intelligence and biotechnology present ethical challenges. And, in an increasingly pluralistic culture, an agreement about how to achieve the common good can no longer be counted on. One danger here is that a cautious optimism about the benefits of future technology is overtaken by 'techno-enthusiasm', which does not recognise what Christian faith takes to be foundational truths about humanity. In particular, the assumptions behind the 'transhumanist' vision of enhancing human beings beyond the bounds of what has previously been thought possible, are at odds with Christian theological convictions about what it means to be human.

Welcome to the age of 'surveillance capitalism', an era where our correspondence is filtered through the servers of Microsoft or Google; where our internet searches and purchases are carefully logged and fed back to us as targeted ads; where invisible code is used to tag our online wanderings; and where, in some parts of the world, citizens are routinely tracked using surveillance cameras and facial recognition.

And welcome to the age of large language models and image generation, where free online tools such as ChatGPT can write an undergraduate essay on any topic; where a text

description can be turned into a new artwork or video; and where created-from-scratch human faces (and naked bodies) can be indistinguishable from real people.

In sum: welcome to the age of a revolution in computing that we call artificial intelligence! And welcome to the age of techno-enthusiasm—a time when Silicon Valley billionaires are overwhelmingly upbeat (along with the required show of concern) about where the overlapping domains of AI and bioengineering will take us.

The AI revolution

The ancient quest for longer life is a bizarre history that includes ancient meditative techniques, the search for legendary fountains, alchemy, and even xenotransplantation (monkey testicles). More recent attempts are no less eccentric and include stem cell transplants, hyperbaric oxygen chambers, young blood transfusions, ketogenic diets, nootropic smart drugs, and pulsed electromagnetic field therapy. Though most of these techniques have yet to produce any real results, life expectancy *has* doubled over the last century in the developed world, largely through improvements in public health and medicine, economic growth, and developments in the areas of nutrition, behavior, and education.[1] North Americans born in 2000 for instance, can on average expect to live to nearly 80 years, compared to just 47 years in 1900.[2]

Despite the artificial intelligence tidal wave of 2023, which brought AI awareness to all, machine 'intelligence' has been with us for decades. Computers have been beating chess grandmasters for over thirty years, almost-autonomous vehicles have navigated traffic for almost as long, and stock market buy/sell 'decisions' have also been made by machines since the last century.

However, with increased computing power and fuelled by massive amounts of 'training data' (and electricity—but that's another story), machines that are designed to mimic the 'neural networks' of the human brain now offer text and visual outputs that are often impossible to distinguish from the work of human creators.

AI offers many beneficial possibilities in a range of areas. Commonly cited benefits include those in *healthcare* (e.g., diagnoses, drug design, patient management and treatment plans); *business* (enhancing productivity by automating tasks and data analysis, fraud detection, customer service); *industry* (robot manufacturing, quality control), *education* (individualised tutoring, document editing, automated assessment), *personal life management* (smart assistants, schedule management, fitness tracking); and *research* (analysing vast data sets, predictive modelling). In short, the useful applications of the power of 'intelligent' computing for social good and efficiency are claimed to be myriad.

However, while the possible benefits of harnessing the power of machine learning range from the obvious to the previously unheard of, the ethical challenges are numerous, from familiar problems, such as job losses due to new technology, to theologically profound questions surrounding what it means to be human. Before turning to specifically theological questions, let me briefly enumerate some of the concerns that Christians might share with other perspectives. These are not comprehensive and there are overlaps between them.

The massing of personal information in the hands of those who don't have the welfare of all in mind is a familiar but escalating problem, because AI requires huge data sets for its 'learning'. Those data sets are amassed from emails, locations, smartphone photos, web interactions, physical actions (captured by CCTV), and the swathes of information online. One example: in 2019, the *New York Times* harvested facial images from a public camera feed. They then spent $100 on facial recognition software, and, for ethical reasons, they chose to identify only one person (with his permission): a local academic. They know what he had for lunch, who he met, and what time he came and went.[1]

The accuracy problem: AI outputs are dependent on the 'training' of the system and on the data set that an AI is trained on. Its outputs are only as good as the information it imbibes, so biases or factual errors in the training data can be reproduced in the AI's results. In 2015, after identifying a bias against women in its recruiting tool driven by AI, Amazon attempted unsuccessfully to retrain the tool. In 2018, the tool was deemed unusable because the AI continued to sort applicants in ways that would be deemed discriminatory.[2] While adept at

finding patterns, AI doesn't understand concepts such as objectivity or truth; its outputs are statistically probable combinations, with no necessary correlation to evidence or truth value.

The alignment problem: Early iterations of AI chatbots were turned off within hours after they 'went rogue', emulating the worst of the online human interactions they had been trained on.[3] Just as AI systems are not 'truth aware' neither can they be morally discerning. If trained on imbalanced or unfiltered data, AI can amplify existing biases. The alignment problem is about how to ensure that AI systems are aligned with human values and principles.

The autonomy problem, associated with the alignment problem, occurs when AI is given control of 'decision making' in the interests of efficiency. How is it possible to ensure that AI controllers of systems such as power grids, telecoms or stock markets make the 'right' decisions? Self-driving cars exemplify the challenge: on the one hand, there are clear benefits to AI controlling a car by reacting more quickly and being more 'alert' than a human driver. On the other hand, AI makes 'decisions' according to built-in parameters and statistical likelihoods but is impervious to moral responses that a human makes in a split second—for example, swerving to avoid a pedestrian or kangaroo even if it means damaging the car or endangering another passenger. This raises critical questions: Who is responsible for an accident caused by an AI? Can pre-programming truly encompass every moral dilemma a car might face on the road?

The black box problem: the neural networks on which AI is based are by their nature opaque and impervious to close inspection. Unlike traditional computer programs, where every step can be followed, it is virtually impossible to analyse exactly why an AI makes a particular 'decision' because the decision-making is embedded in literally millions of statistical calculations. This makes it impossible to guarantee outputs and raises ethical concerns about whether AI can be 'trusted' in high-stakes applications.

Evidently, there are significant ethical questions to be asked about the use of this sort of technology. These ethical concerns arising from AI are going to be shared by many people and are not concerns specific to Christian faith. However, while the processes and products

of science and technology have always raised ethical questions, in an increasingly pluralistic culture an agreement about how to achieve the common good can no longer be counted on.

Enhancing human bodies

In considering the vision of what AI and bioengineering might mean for enhancing human beings, Christian theological concerns come to the fore. In particular, how should we think about the use of AI and biotechnology to enhance human bodies beyond the bounds of what has previously been thought possible? This is the realm of what is commonly called transhumanism.

In a nutshell, the transhumanist vision is to pick up the process of biological evolution and keep it rolling forward faster. Humans have evolved to this point, transhumanists say, but we have so far to go. Integrating AI and biotechnology puts it within the power of humans to take charge of our own evolution. In a move that takes humankind from therapy to enhancement, some techno-enthusiasts, including the World Transhumanist Association, say that we don't just want to be well; 'we want people to be better than well'.[4] This is a vision of endless improvement and amounts to a new ideology: a 'secular version of immortality on its own terms',[5] as one commentator has said. Are there strains here of Genesis 11 and a new Tower of Babel?

In the transhumanist vision, human functioning as we know it is not enough; this vision looks to the use of AI and biotechnology to *enhance* human physical and mental capacities beyond this. It is about prolonging and 'improving' human life. Some, such as Elon Musk and Bryan Johnson, are investing in brain-machine direct connections in order to enhance human cognitive ability.[6] Or listen to one enthusiastic supporter, Ray Kurzweil, once a senior engineer at Google, who looks forward to

> *computers having human intelligence, our putting them inside our brains, connecting them to the cloud, expanding who we are. Today, it's here, in part, and it's going to accelerate. We're going to get more neocortex, we're going to be funnier, we're going to be better at music. We're going to be sexier. We're really going to exemplify all the things that we value in humans to a greater degree.*[7]

However, a symbiosis between humans and machines is not brand new and it is certainly not all bad. In fact, a deeply committed Christian man from Melbourne created the world's first bionic recreation of a human sense. Graeme Clark, at the University of Melbourne, led the team that developed the cochlear implant, and now scientists and engineers at the same university are working on a bionic eye. As well, they have inserted an electrode into a patient's brain allowing them to control a computer just by thinking.[8] Long-term this technique will allow a person to control bionic limbs. David Grayden, who is the Clifford Chair of Neural Engineering at Melbourne, is also a Christian. David is happy to say that he is following in the footsteps of the master, enabling the deaf to hear, the blind to see and the lame to walk. These people believe that God has given humanity the calling and often the means to restore people to health and, what we might loosely call, normal human functioning.

But the ideology of the transhuman vision goes further. For the transhumanist, human functioning as we know it is not enough; the point is to use technology to *enhance* human physical and mental capacities beyond current human functioning. In short, it is about prolonging and 'improving' human life. For the techno-enthusiast, as well as being faster, stronger, sexier, more intelligent, the holy grail is the dream of life without limit.

The rudiments of the medical technology necessary to continuously repair the body already exists. The CRISPR gene editing technique has been used to inactivate a gene in mice that causes cellular senescence—the old age of cells where they no longer reproduce.[9] The ultimate end of such research is to ensure that cells continuously rejuvenate themselves. Meanwhile 3D printers are already used to build up tissues layer by layer.[10] Who knows how long it will be before medicos can order a 3D printed kidney made to a patient's own biological specifications?

Further into the realms of science fiction are transhumanist hopes that one day it will be possible to map and download all the information stored in a human brain. Yes, that raises certain questions about human consciousness, but for those who dream of such things, the logic goes like this: If the essence of a human being is the information contained in the brain,

and if the body can be separated from that information, then the information that makes you *you* could be downloaded to a hard drive. And then, surely, it could be uploaded to the cloud. So, *you* would be there, stored somewhere in cyberspace in your own personal 'I' cloud.

Techno-redemption

Before considering some theological questions about where AI and biotechnology might lead, let me point to one more, very serious, proposal by some philosophers—that of 'techno-redemption'. Some people argue that genetic engineering has brought us to the place where we should take seriously the possibility of morally enhancing future humans. While the word 'sin' is not used, the proponents argue that there are things very wrong with the world, and indeed very wrong with many human beings.

In a networked world, where the internet allows an individual access to global companies and government databases, a single person has unprecedented destructive powers. One individual can hack a telecommunications or health insurance database, and a rogue research scientist could release biological havoc on the world. Humanity is increasingly at the mercy of individuals who are capable of doing great harm to society and the planet.

This is the basis of the argument by philosophers such as Australian, Julian Savulescu, at Oxford University, who say that we should consider genetically modifying future humans to reduce the risk that people will engage in acts of what they call ultimate harm.[11] Here is one commentator's paraphrase of the argument:

> *Given that technology will increasingly give all of us the power to inflict Ultimate Harm — either quickly and individually in the case of bioterrorism or slowly and collectively in the case of climate change — what needs to change is us. If the world will blow up if just one of us pushes the self-destruct button — or if all of us won't stop pushing the climate change button — than what we need are human beings who can be trusted not to push that button. What we need are better people.*[12]

However, 'What we need are better people' begs the crucial question: who defines 'better'? This argument, and the techno-enthusiast enterprise, depends on a vision of the good: a clear vision of what it is to be moral, what it is to be a better person. Who will decide these things in contexts where no moral framework is agreed?

When I first read C. S. Lewis's essay, *The Abolition of Man*, I was struck by its prophetic nature. Written 80 years ago, it paints a dystopian picture of a time when a technologically powerful society is no longer constrained by an agreed set of moral principles. If there is no overall moral framework, then we are all subject to our own wills and desires and natural impulses, in turn shaped by our genes, homelife, community, culture and more. Going down the road of humanity taking control of its own evolution, says Lewis, will eliminate what makes us most essentially human; the result is the abolition of humanity. And, even as ethicists argue, those who seek power or wealth will get on with using technology for their own interests. No longer driven by a clear moral framework, human beings are driven by their own wants and desires rooted in unconstrained human nature. Let me quote from Lewis's book (edited for gender neutrality):

> *If any one age really attains, by eugenics and scientific education, the power to make its descendants what it pleases, all people who live after it are the patients of that power. … The final stage is come when humanity, by eugenics, by pre-natal conditioning, … has obtained full control over itself. … [But] the power of humanity to make itself what it pleases means the power of some people to make other people what they please.*
>
> *At the moment, then, of humanity's victory over nature, we find the whole human race subjected to some individuals, and those individuals subjected to that in themselves which is purely 'natural' – to their irrational impulses. … Humanity's conquest of Nature turns out … to be Nature's conquest of humanity.*[13]

Theological concerns

Turning to theological questions that challenge techno-enthusiast visions of a human/AI-created utopia, the overarching question raised by future technology is about the nature of humanity. What is a human being? In a world where all things seem fluid, the Christian faith holds to convictions that directly oppose techno-enthusiast aspirations. Before exploring these, let me first flag two ancient heresies about which the Christian church has always been wary.

One is Pelagianism, the core of which is the idea that humans are essentially good; wrongdoing arises from bad habits, rather than being an intrinsic aspect of who we are as self-centred beings. If we try hard enough, we can, by God's grace, live a truly moral life for a day or week or month. Outside Christian circles, this assumption of essential human goodness leads to a great confidence in human progress; technology gives us the tools and humans can be trusted to use them for good.

The other ancient heresy is Gnosticism, which sees the world so 'spiritually' that the material world (including the body) is seen as holding us back and of no intrinsic value. The immaterial soul or mind is the permanent self, and its wellbeing is fostered by knowledge (gnosis), which is the highest goal of life. Ignorance rather than sin is the human problem; the remedy is to attain spiritual knowledge. In the words of one critic, 'transhumanism in all its hubris … relies on a retooled Gnostic heresy that is very prevalent today … and can therefore be divided with no consequences.'[5]

With these boundaries in mind, six fundamental questions arise that go to the heart of what biblical faith teaches about what it means to be human.

Firstly, *what is our theology of the common good?* If you ask a good Presbyterian what it means to be human, one answer you might get is that the chief purpose of human beings is 'to glorify God and enjoy him forever' (from the Westminster Shorter Catechism). If you ask a Catholic, they might say 'to know, love, and serve God' (from the Catechism of the Catholic

Church). And if you ask Jesus, he would add that such delight in God is seen in how we love our neighbour as ourselves (Mark 12:31, citing Leviticus 19:18). Christians thus have a clear understanding of our responsibility to serve not ourselves but the common good. Life is not about technological self-actualisation. The affirmation of our embodied worth, exemplified in a life of self-sacrifice, which is shown most clearly in the life of the one who died for us while we were yet sinners, is at the heart of Christian faith.

Talking of sinners, a second question is this: *What is our theology of sin?* In Christian understanding, in women and men – though made in the image of God - sin runs deep. Freedom from its grip entails more than tweaking genetics—if that were possible—so that someone does not engage in fits of rage or antisocial tendencies. Sin runs much deeper than what techno-enthusiasts might want to excise from the human genome. Further, in Christian understanding, sin distorts the human race: it has corporate as well as personal dimensions. Only through the costly work of God in Jesus Christ can it ultimately be healed. The Spirit of God works ceaselessly to bring this about, but we are not whole yet. To have faith that technology, informed by the ethics of an elite group, could purify the human race, is, from a Christian point of view, both absurd and dangerous.

This leads to another question for Christians: *What is our theology of redemption or liberation?* For the techno-enthusiasts I have described, freedom from sin and evil takes place within this material non-transcendent world through technology offering solutions to all problems; 'heaven' is what humans make of it. However, Christians look to *God* to ultimately transform all things. Christians look for the transformation of the 'groaning' creation that the Apostle Paul reflects on in Romans chapter 8. Christians do not live under the delusion that humans can rescue themselves. Christian hope is in God, the creator and liberator, to bring in that new creation, which the Bible calls 'the new heavens and earth'. Christian hope is in the saving work of the Lord Jesus Christ to deliver us from the consequences of sin and the brokenness in our earthly lives. In the short term, Christian hope is in the work of the Holy Spirit to do this in us day by day, remaking us to be more and more like Christ. The Christian hope is not in genetics to reverse the story of Genesis chapter 3, but in God's liberating transformation of all there is, guiding human history towards its true end, the common good of creation as a whole.

A fourth question: *What is our theology of the body?* Christian hope is found in the resurrection of the body, whatever that looks like. Since we will be embodied in the new creation, our present bodies matter; they are more than a collection of X terabytes of information loaded into our flesh and bones that could, hypothetically, be downloaded to a robot body or uploaded to the cloud. Paul makes it clear in 1 Corinthians 15, as do the varied Gospel accounts of the resurrection, that human bodies are the base and core of who each of us is. Further, 'the Word took on our flesh and dwelled among us,' in the prologue to the Gospel of John, is the ultimate Christian affirmation of our physical embodiment. Christians cannot go along with a view that says our essential self could be saved as data on a hard disk. Nor can Christians imagine that any amount of data and AI processing power, which is not integral to its own unique body, could 'become human'.

Speaking of the frailties of the body raises the question of suffering. *What is our theology of suffering?* Christian faith reflects a mixed attitude to suffering. On the one hand, suffering is a universal human experience, seen as rooted in our brokenness due to sin. The good news of liberation from sin means that it is our duty to ameliorate suffering. However, Christians also find meaning in suffering. Suffering can be redemptive and purposeful, especially as we see ourselves united with Christ's own suffering on the cross. Suffering can be an opportunity for spiritual growth, for empathy and acts of compassion, and for deepening faith through trials. Suffering also highlights the hope of ultimate restoration, when 'death will be no more, nor mourning and crying and pain' (Revelation 21:4). The aspiration to eliminate all suffering by technological wizardry, sits most uncomfortably with Christian theology.

Finally, in contrast to the extremes of techno-enthusiast aspirations, *what is our theology of life everlasting?* Christians await the new heavens and the new earth, inaugurated in God's time and not in ours. But this 'life of the ages' (as the New Testament puts 'eternal life') begins in this age (cf. John 3:16). So, Christians embrace the responsibility to be creative, to address suffering, to restore human functioning, even to prolong life—but not at all costs. Life is not prolonged by Christians in the interests of fleeing from death, which brings the hope of resurrection and our ultimate joyful encounter with God face to face.

Where does all this leave the Christian in an AI world? Perhaps it leads us to offer two cheers for techno-enthusiast idealists. Yes, things could be better; yes, there are things wrong with the world and humanity. Perhaps the enthusiastic vanguard of AI and biotechnology, looking for a better world, is not so much anti-God as misguided. And, as Christians, we do look forward to renewed bodies and minds—enhanced bodies and minds, way beyond human functioning as we know it now. But that is the gracious, loving work of God, brought into our present by the Spirit of Christ—the new creation which we look forward to in faith.

Chris Mulherin is the Executive Director of ISCAST–Christianity and Science in Conversation, and he teaches philosophy at the University of Divinity in Melbourne. He is the author of *Science and Christianity: Understanding the Conflict Myth.*

References

1. Chinoy S. We built an 'unbelievable' (but legal) facial recognition machine. *New York Times*. https://www.nytimes.com/interactive/2019/04/16/opinion/facial-recognition-new-york-city.html (April 16, 2019, accessed 19 July 2024).

2. Dastin J. Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/ (October 11, 2018, accessed19 July 2024).

3. Kraft A. Microsoft shuts down AI chatbot after it turned into a Nazi. *CBS News*. https://www.cbsnews.com/news/microsoft-shuts-down-ai-chatbot-after-it-turned-into-racist-nazi/ (March 25, 2016, accessed 19 July 2024).

4. Humanity+. Mission, https://www.humanityplus.org/about (2024, accessed 19 July 2024).

5. Showalter B. Google director's push for computers inside human brains is 'anti-Christ,' 'Human Rights Abuse,' theologians explain. *Christian Post.* https://www.christianpost.com/news/google-directors-hope-for-computers-inside-human-brains-is-anti-christ-human-rights-abuse-theologians-explain.html (March 17, 2017, accessed 19 July 2024).

6. Marsh S. Neurotechnology, Elon Musk and the goal of human enhancement. *The Guardian.* https://www.theguardian.com/technology/2018/jan/01/elon-musk-neurotechnology-human-enhancement-brain-computer-interfaces (January 1, 2018, accessed 19 July 2024).

7. Sulleyman A. The singularity: AI will make humans sexier and funnier, says Google expert. *The Independent*. https://www.independent.co.uk/tech/singularity-artificial-

intelligence-humans-sexy-funny-ai-music-art-google-futurist-engineering-ray-kurzweil-a7633481.html (March 16, 2017, accessed 19 July 2024).

8. University of Melbourne. Helping people with paralysis operate computers and phones via thought. https://www.unimelb.edu.au/alumni/impact/research/helping-people-with-paralysis-operate-computers-and-phones-via-thought (2022, accessed 19 July 2024).

9. Wang W, Zheng Y, Sun S, Li W, Song M, Ji Q, Wu Z, Liu Z, Fan Y, Liu F, Li J. A genome-wide CRISPR-based screen identifies KAT7 as a driver of cellular senescence. *Science Translational Medicine*. 2021 Jan 6; 13(575): eabd2655. https://pubmed.ncbi.nlm.nih.gov/33408182/ (accessed 19 July 2024).

10. Barber C. 3D-printed organs may soon be a reality. 'Looking ahead, we'll not need donor hearts'. *Fortune*. https://fortune.com/well/2023/02/15/3d-printed-organs-may-soon-be-a-reality/ (February 16, 2023, accessed 19 July 2024).

11. TED Guest Author. As a species, we have a moral obligation to enhance ourselves. *TED.com.* https://ideas.ted.com/the-ethics-of-genetically-enhanced-monkey-slaves/ (February 19, 2014, accessed 19 July 2024).

12. Walsh B. The case for genetically engineering ethical humans. *OneZero.* https://onezero.medium.com/the-case-for-genetically-engineering-ethical-humans-b44c17b9e3d6 (July 27, 2018, accessed 19 July 2024).

13. Lewis CS. *The abolition of man.* Glasgow: Collins Fount Paperbacks; 1981, p.36-41.