ARTIFICIAL INTELLIGENCE

A Theological Approach

Calum Samuelson

THE BIBLE PROVIDES invaluable insight regarding who we are as humans. Throughout history, human culture and knowledge have taken many forms and have often progressed; but human nature itself has not changed. Technological innovation can sometimes obscure this truth, especially when it enables humans to accomplish unprecedented feats of power, skill or foresight. However, unless we understand our purpose, weaknesses and trajectory as humans, even the most advanced technologies will simply make us more efficient in repeating the same mistakes we have always made.

This is especially true of artificial intelligence (AI)—which is perhaps the most pressing case in point. Whether from news stories or fiction, most people have by now become aware of AI's purported potential to revolutionise human existence. Indeed, some insist that such a change is already well under way. Cars can drive themselves, robots can perform intricate surgeries, and computers can convincingly converse with people. But what exactly distinguishes AI itself from cars, robots and computers?

At the most basic level, AI refers to the ability of computer systems to perform tasks normally associated only with human intelligence. Yet because this distinction is continually shifting, a better way to understand AI is that it involves computer systems improving their own performance independent of human intervention—otherwise known as 'machine learning'. While there is little doubt that AI will bring about significant changes to our world, it is more difficult to pinpoint what its main role will be. The most prevalent predictions can be generalised into three scenarios:

This article is an extract (slightly adapted) from a Jubilee Centre publication (Cambridge), Calum Samuelson, *Artificially Intelligent? Grappling with the Myths, Present Realities and Future Trajectories of AI* (2018), available at https://www.jubilee-centre.org/ebooks/artificially-intelligent-grappling-with-the-myths-present-realities-and-future-trajectories-of-ai, accessed 9 June 2020.

The Way, 59/3 (July 2020), 41–50

1. AI will eventually make its biggest impact as a completely autonomous and perhaps conscious entity which overtakes humanity.
2. AI will prove to be most significant when it becomes seamlessly integrated with biological human bodies, vastly increasing abilities and prolonging life.
3. AI will continue to be harnessed by humans as all technological innovations in the past have been—a morally neutral tool that can be used to accomplish tasks with greater ease and efficiency.

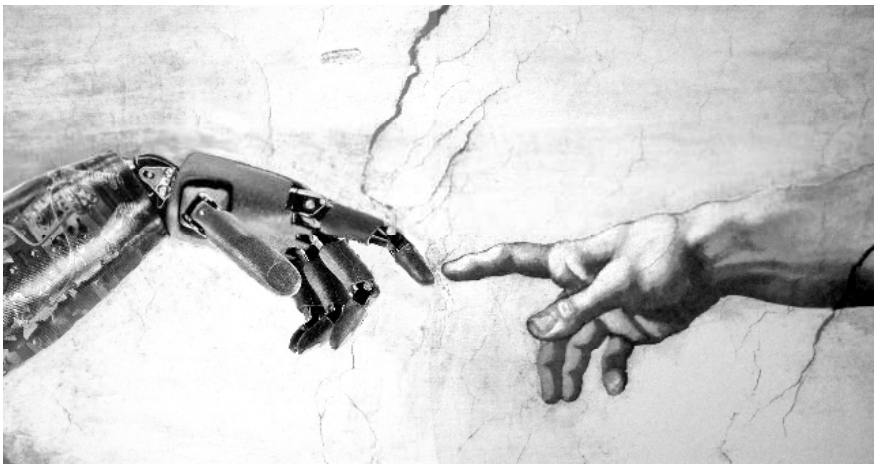
Although the argument cannot be made here, I am convinced that the final scenario is correct and take it for granted in this essay—AI is a *tool that amplifies human nature and behaviours* rather than transforming them.[1] Accordingly, I shall address AI obliquely by concentrating on human nature particularly, focusing on three biblical themes: the *Imago Dei*, the Fall and eschatology. In particular, the *Imago Dei* helps us identify which human qualities and characteristics AI should seek to facilitate or enhance. The doctrine of the Fall helps us diagnose how human imperfections and malevolence influence the development and application of AI. Finally, biblical eschatology helps us anticipate, imagine and yearn for our ultimate destination and think critically about different AI-powered futures.

Imago Dei

Exploring the best dimensions of humanity is essential to the task of using AI tools to amplify good and promote human flourishing. AI experts are eager to parse the distinction between humans and computers. In this effort, it is common to invoke illustrious human feats such as Michelangelo's paintings, Bach's cantatas or Einstein's theory of relativity. This method of distinguishing between human and AI is unsatisfactory, not least because it neglects the majority of people who have ever lived. Most importantly, however, such cursory assessments of humanity's greatness fail because they measure accomplishments divorced from the role of purpose. The *Imago Dei* helps us understand human purpose better.

The belief that humans are made in the image of God has rightly occupied a central position in the Christian consideration of AI to date. Creativity, reason and morality have largely dominated as the primary

[1] This argument is spelled out in sections 1 and 2 of *Artificially Intelligent?*



dimensions of the *Imago Dei* in the last millennium.[2] Despite their importance, these dimensions may have had more popular currency in the modern period (when there was more social consistency and structure) than they do today in the fragmented and pluralistic world where AI is making its mark. Indeed, considering the perilous state of human identity in the postmodern era, it is hardly coincidental that humans are increasingly compared to computers. Consequently, this section considers the *Imago Dei* through the lens of *relationships, responsibility and self-giving love*.

According to the Bible, humans are explicitly created in the image of a *relational* God, the implication being that we are only fully human when in meaningful *relationship* with others (Genesis 1:26). The metaphor of the Christian community as a body teaches that every member plays an integral role (1 Corinthians 12:12–31). Also, the fruit of the Spirit is always manifested in relational contexts (Galatians 5:22–23). This raises important questions for the development of AI tools.

Moreover there can be no single version of the 'ideal human' because each possesses different qualities and gifts in varying degrees and arrangements, and each exists in relationship with others. Consequently, some are suggesting that it is better to design a range of AI tools to do different tasks rather than a single tool that mimics humans completely.[3] This also has important implications for current discussions about 'digital

[2] Roughly speaking, one can recognise that creativity was explored in the Renaissance, reason in the scientific revolution and morality in the Enlightenment.

[3] See G. Andrew D. Briggs and Dawid Potgieter, 'Machine Learning and the Questions It Raises', in *From Matter to Life: Information and Causality*, edited by Sara Imari Walker, Paul C. W. Davies and George F. R. Ellis (Cambridge: CUP, 2017), 468–486, here 478.

personhood' and 'digital subjects', since these terms suggest it is possible to know someone apart from a relational context.[4] Finally, the fact that humans are created for relationships can help explain the tendency to anthropomorphize technologies such as AI, and can also shed light on the human vulnerability to computers that simulate humans by design.

Another implication of humans being created in the image of God is that they have *responsibility*. God is the supreme, faithful sustainer of all Creation but has also entrusted humans with the unique responsibility of caring for and ruling over that Creation. Much more than a mere task or goal, this responsibility requires the entire human being to act like an 'angled mirror' which simultaneously reflects the lordship of God to Creation and the praise of all Creation back to God.[5] The importance of responsibility in the realm of AI may be the most needed element.

Several Christians developing AI today understand their work as a clear example of subduing the earth (Genesis 1: 28); others think of it more as an aspect of serving people and society with love.[6] In either case, there is a clear difference between designing AI tools to aid in the responsibility of wisely ruling Creation and designing them to rule so that humans can shirk the weight of that responsibility. Already, one can perceive small ways in which humans are abdicating their responsibility for ruling through AI, whether by using autonomous weapons, foetus screening, employee profiling or criminal facial recognition. Increasingly, if the AI says a decision is right, the human users will execute it. This is not only the definition of *irresponsibility*, it also suppresses human creativity by assuming that difficult ethical decisions can be avoided or even eliminated.[7]

The *self-giving love* of God—which has always existed in the dynamic relations of the Trinity—flowed outward in the act of creation and was eternally enacted in the sacrifice of Jesus on the cross. Because they are created in this God's image, one characteristic of humans is the capacity to love in a manner that considers the needs of others above self. Culture at large often only praises this type of love if it takes the form of heroism, as it does in Hollywood—even the deeply shameful crucifixion

[4] See Olga Goriunova, 'The Digital Subject: People as Data as Persons', *Theory, Culture and Society*, 36/6 (November 2019).

[5] Tom Wright, *The Day the Revolution Began* (London: SPCK, 2016), 100.

[6] See *Artificially Intelligent?* appendix, 37.

[7] Whilst advocating a robust understanding of (and emphasis on) *human* responsibility, I do not address the particulars of *how* ethical decision-making algorithms should be constructed here because this will vary across cultures, contexts and organizations. There is already a large body of interdisciplinary work being done in this area, particularly around guidelines for autonomous weapons and self-driving cars.

has been turned into a grand act of courage and fortitude. Others dismiss the value of sacrifice in favour of more empirical, scientific accomplishments, to the extent of marginalising God altogether. One influential statement about AI claims, 'everything that civilisation has to offer is a product of human intelligence'.[8]

Whatever the world says about love, Jesus claimed that there is no greater love than to lay down one's life for one's friends (John 15:13). Within this logic, the widow who gave two mites was praised above the lavish tithers (Luke 21:1–4), the quiet tax collector was the one made right with God (Luke 18:9–14) and small children are singled out as possessors of God's Kingdom (Matthew 19:13–14). Human greatness is not found in mighty, memorialised achievements, but in seemingly 'small', self-giving acts of devotion, humility and sacrifice.

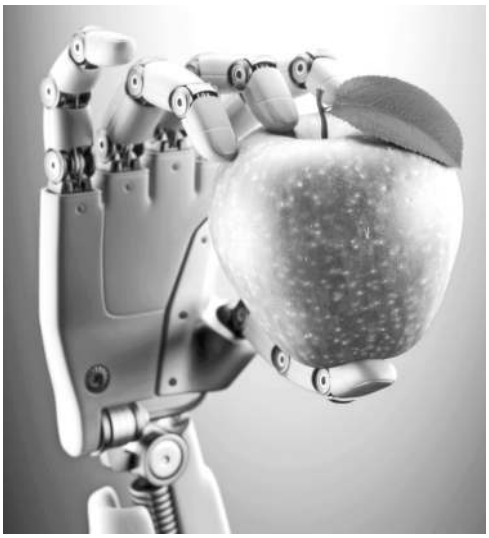
As AI becomes increasingly common, it is important to remember that love always prioritises the other. AI may help people feel happier, be more efficient, obtain more knowledge and even feel more ethical, but if it does not improve human relationships it is ultimately misdirected. Accordingly, people should be very cautious about seeking to outsource or automate the most common and apparently mundane manners in which they give themselves in love to others. The simple gift of listening is rapidly being replaced by AI. The command to weep with those weep (Romans 12:15) is being threatened by AI tools that detect our mood and tell us how to fix it. The practice of hospitality in which we open up our homes is being superseded by virtual interactions. If *agape* love were simply another task that required energy to perform, then it would make sense to continue designing AI tools that preserve energy. But *agape* love is not a separate task to be performed at the end of the day like some leisure activity. It is something that must be *practised and developed*, and often the best way to do this is by washing the feet that no one else wants to wash (John 13:1–17).

Love always prioritises the other

The Doctrine of the Fall

Whereas the previous section explored the goodness of humanity in Creation, this one highlights its shortcomings. A major concern for Christians who are developing AI is that secular thinking is not equipped to account adequately for or to anticipate the realities of imperfection and malevolence in human nature and the world. Of course, secular

[8] Stephen Hawking, *Brief Answers to the Big Questions* (London: John Murray, 2018), 183.



programmers and developers acknowledge that glitches can plague a computer system and that sometimes criminals hijack a piece of good technology for a bad purpose. But, for the most part, AI development buzzes with an optimism that believes sustained effort and education can eventually help humanity overcome all its problems and perversions.

The Bible sees things differently. God created a world that was 'very good', but it has fallen from that status because of sin—which is anything that obstructs relationship with God. Humankind can naturally recognise entropy, atrophy, disease, corruption and brokenness of all kinds as deviations from an ideal situation, but can also become tragically resigned to the idea that these things are simply woven into the fundamental fabric of the universe. Crucially, the doctrine of the Fall helps make sense of the tension between desired behaviour and actual behaviour, and helps Christians consider how this tension might influence the development and deployment of AI. In order to do this, it is necessary to examine both the *depth* and *breadth* of sin. The *depth* of sin reaches to the very core of our being and cannot be encompassed by a binary system of 'rights' and 'wrongs'. Jesus taught that even perfect 'right' actions can be sinful if done with the wrong posture of heart. It may be possible to distance oneself from particular external sins, but no one is ever far away from the allure of pride and self-assurance.

Applied to AI, this truth has two major implications. First, it means that attempts to transcend human faults and discover 'perfect morality' using AI are misguided. Indeed, perfection should never be attributed to machines, created as they are by imperfect humans. Second, it means that even the best AI developments can have negative consequences. Some of these are caused by glitches or programmer bias. More insidious, however, are AI applications that seem supremely good or helpful but ultimately turn hearts away from God (such as a financial tool that ends up increasing greed or a voice replication tool that ends up enabling deception). Some people fear enslavement to AI through *oppression*, but

we are already becoming enslaved through the subtler route of *obsession*.[9] It may well be that efficiency and knowledge will be the predominant idols of the AI age.[10]

In addition to the *depth* of sin, its *pervasiveness* touches every corner of the world. In a hyper-individualistic age, it is easy to interpret passages such as Romans 3:23 ('all have sinned and fall short of the glory of God') as an indictment against particular, personal failures. But Paul's message carries a sweeping universality from which nothing can hide. *All* of Creation groans for redemption as the effects of sin are felt. This means that sin is encountered both internally and externally, individually and structurally.

Just as an excellent policy or strategy can be thwarted by external factors, so also can AI fail owing to user error, corrupt data or false information. It is conceivable that one party, nation or culture could develop a genuinely productive framework for engagement with AI, only to have it disrupted or destroyed by a broken, sinful mindset or system. The pervasiveness of sin must also be considered in a diachronic sense. One of the great falsehoods connected with modern myths of inexorable progress (whether capitalistic, Neo-Darwinian or even 'exponential') is the idea that human morality itself can continually improve.[11] While it is obvious that most humans in the West no longer pillage, rape, burn, imprison or torture other people, one need not look far to uncover modern equivalents in the form of embezzlement, habitat destruction, child abuse, debt slavery and animal cruelty. Humanity does not get 'better' intrinsically, we simply get 'better' at devising ways to justify our crooked actions.

At its most basic level, the pervasiveness of sin confronts the field of AI development, in which progress, success, benevolence and good behaviour are simply taken for granted. One must not only consider the impact of individual sin, but also of sin within every other person and institution with which an individual interacts. One direct implication for AI development could be to design systems in a way that *expects them as a rule* to break down, to be misused and to affect unexpected stakeholders.[12]

[9] There are valid concerns around AI's potential use for oppression, especially in authoritarian states. The AI-assisted Social Credit Register introduced in China is perhaps the best example of an effective surveillance and social control mechanism, see https://www.wired.co.uk/article/china-social-credit. George Orwell's *Nineteen Eighty-Four* and Aldous Huxley's *Brave New World*, chart these trajectories vividly.

[10] Crucially, idolatry or obsession facilitated by AI tools can grow much faster than other types since AI increases efficiency by definition.

[11] For exponential progress, see Ray Kurzweil, *The Age of Spiritual Machines* (London: Penguin, 2000).

[12] See Parliamentary Select Committee for Digital, Culture, Media and Sport, interim report, 'Disinformation and "Fake News"', available at https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/36302.htm, nn. 67–73 on the misuse of AI and nn. 243–248 for unexpected impact.

If the world was not called 'perfect' even before the Fall, we should hardly expect that we can make it perfect through AI now.[13]

Eschatology

In addition to appreciating humanity's purpose and sinfulness, a holistic view requires comprehension about humanity's trajectory and ultimate destination. According to the Bible, this trajectory is inseparable from the redemption already inaugurated in the person of Jesus and headed towards a supremely good New Creation after the end of this age. One could even argue that the only type of true inexorable growth that is possible in the universe is growth in Christlikeness, which by the Spirit's power will continue for all eternity. Regardless of whether a linear or cyclical view of time is espoused, it is not uncommon for humankind to yearn for an ultimate destination beyond time, and many generations have thought the world will end with them. It should be no surprise that much of AI dialogue also yearns for a different future and ultimate end for humanity. Hence there is the need to highlight the gravity of long-term effects and the need for goal-orientated trajectories of AI, and one of the best ways to do this is to consider what the Bible says about humanity's ultimate end.

First, Paul clearly teaches that resurrected humans will not be spirits without bodies (1 Corinthians 15). This has important implications for various agendas which view the human body as a disposable inconvenience and hope that AI will help humans eventually to discard it. Second, eschatological pictures in the Bible envision the flourishing of non-human Creation. This point has vital implications for the care of animals and the environment, for it seems that ultimate symbiosis with the New Creation is meant to be an outflowing of human interaction with Creation in this life.[14] Third, the Bible portrays a dynamic pan-ethnic relational community existing in the New Creation. This challenges aspirations positing seamless technological uniformity, compatibility or even complete 'monism'.[15] Last, the Bible emphasizes the importance of simplicity and purity (Ephesians 5:26–27) in the Kingdom of God, which belongs to

[13] Orthodox Christians have long understood Adam and Eve more as innocent children than perfect humans. Additionally, some would point out that the Garden could not have been perfect if it contained a deceptive serpent and a tree containing the knowledge of evil.

[14] The traditional reading of 2 Peter 3: 10 has tended to emphasize the destructive nature of the fire, but several scholars are trying to recover the true reading as a 'refining fire'. See Richard Middleton, *New Heaven and New Earth: Reclaiming Biblical Eschatology* (Grand Rapids: Baker Academic, 2014), 160–163.

[15] See Jan H. Naude, 'Technological Singularity and Transcendental Monism: Co-Producers of Sustainable Alternative Futures', *Journal of Futures Studies*, 13/3 (2009), 49–58.

the little children (Matthew 19:13–14) whose play characterizes and energizes the perfect peace we will know there (Zechariah 8:3–5). This is quite distinct from some secular narratives which aspire to vast knowledge, efficiency and complexity.

While Christians can be confident in Christ's return, humility and vigilance are the best postures for discussing *how* this will come about. Just as transhumanists and technologists have overestimated the pace at which 'exponential' AI development will usher in a New Age, so some Christians have also overanticipated the full arrival of the Kingdom of God. It is interesting that a (minority) stream of fatalism regarding environmental destruction also runs through both groups. Some fundamentalist Christians in the USA see the destruction of the environment as a step towards precipitating Christ's return. Some futurists believe that biological life will be superseded since machines can run on solar-generated electricity; for them the ultimate demise of carbon-based life forms is what drives the urgent search for life beyond the need for biological resources.

Will AI help us save the environment and usher in a better age, free from fossil fuels, or will it be the only recourse available after we have destroyed the biosphere? Will Jesus return before or after the planet is hit by a super asteroid? The parable of the wheat and the tares can help Christians navigate the conflicting reports about the world's trajectory, as it insists that both evil and goodwill continue to increase in the world until Jesus' return (Matthew 13:24–30). This means that fear, or naïve optimism, or apathy is not an appropriate mindset, because Christians are called to be alert, joining in the work of the Spirit wherever it may be found.

One practical way to live within this tension is by nurturing a theology of surprise.[16] Rooted in God's often unexpected works of redemption, this way of viewing the world *actively anticipates* God doing surprising things as Christians act as salt and light in the world. A theology of surprise protects against excessive commitment to narrow programmes or agendas, as both God's warnings and God's blessings come in ways that cannot be predicted. With regard to AI, this may mean that Christians encounter real hope in the places with which they are least comfortable and fear in the places where they least expected to find it.

[16] Although not known to have used the specific phrase himself, Lesslie Newbigin frequently promotes this idea and discusses the ways he was surprised by God while working as a missionary in India. See *The Open Secret: An Introduction to the Theology of Mission*, rev. edn (Grand Rapids: Eerdmans, 1995), 61.

Dialogue about what we ultimately yearn and hope for can be powerfully inviting, and some may find that these eschatological topics present a good way of engaging transhumanists and technologists in meaningful conversation. Public dialogue is increasingly turning to questions about what an ideal society *should* look like, and Christians should capitalise on this opportunity by looking forward to what perfect eternity *will* look like. This practice is deeply demanding because it requires the active deployment of our imagination in tandem with the mysterious movements of the Spirit, but for that very reason is also infinitely more valuable than anything Christians do without the help of God.[17]

Stewards of Creation and Citizens of Heaven

Leaders of all types have the obligation to ensure that AI does not simply amplify the current trajectory of present realities such as individualistic capitalism, and it has been argued here that a keen understanding of humanity is crucial for this endeavour. In particular, leaders must take seriously humankind's propensity towards malevolence (doctrine of the Fall) while being rooted in its ultimate calling (*Imago Dei*) and directed towards its final end (New Creation). There can be no doubt that AI will transform the world as we know it. As ambassadors and servants of Christ, Christians especially should strive to direct the impacts of AI in ways that help people live life to the fullest and bless the communities, cities and countries where they live. Just as globalisation, despite its benefits, has accelerated the loss of indigenous languages and cultures, it is conceivable that mass, indiscriminate implementation of AI systems could make humans very good at doing things which are not in their best interests. Ultimately, AI tools should help people regain healthier notions about the purpose of life in general. Recapturing both the art of discipline and a sense of human purpose, people can learn to eschew those effects of AI which produce burnout or laziness in favour of those that help them mature and thrive as stewards of Creation and citizens of heaven.

Calum Samuelson is a pastor, chaplain and researcher who is passionate about the intersection of theology with social dilemmas. He lives in Seattle, Washington, and enjoys spending time in nature with his wife and son.

[17] Tom Wright's modern classic, *Surprised by Hope* (New York: HarperOne, 2008), offers several inspiring ideas about how this may unfold practically in the lives of believers. And see Psalm 127:1.