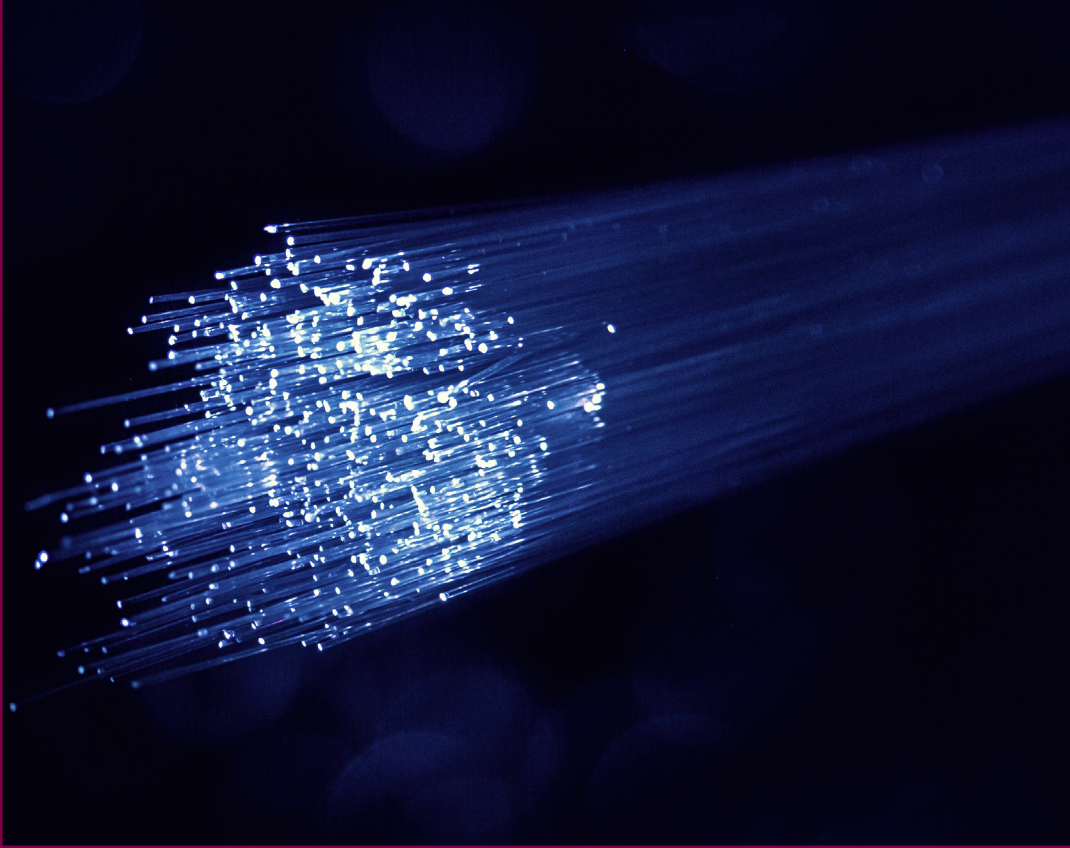


The Challenge of Artificial Intelligence

Responsibly Unlocking the Potential of AI

Enterprise and Faith series

Andrei E. Rogobete

The Centre for Enterprise, Markets and Ethics

The Centre for Enterprise, Markets and Ethics

ENTERPRISE AND FAITH SERIES

The Challenge of Artificial Intelligence

Responsibly Unlocking the Potential of AI

Andrei E. Rogobete

First Edition, 2023

ISBN: 978-1-910666-24-1

Published by:
Centre for Enterprise, Markets and Ethics, 31 Beaumont Street, Oxford OX1 2NP

Design by:
Joy Graphics Ltd,
31 Mays Crescent, Northleach, Gloucestershire GL54 3HR

Printed in the United Kingdom by:
Foremost Print Ltd, Unit 9a, Vantage Business Park,
Bloxham Road, Banbury, Oxfordshire OX16 9UX

The Centre for Enterprise, Markets and Ethics

We are a think tank based in Oxford that seeks to promote an enterprise, market economy built on ethical foundations.

We undertake research on the interface of Christian theology, economics and business.

Our aim is to argue the case for an economy that generates wealth, employment, innovation and enterprise within a framework of calling, integrity, values and ethical behaviour, leading to the transformation of the business enterprise and contributing to the relief of poverty.

We publish a range of material, hold events and conferences, undertake research projects and speak and teach in our areas of concern.

We are independent and a registered charity entirely dependent on donations for our work.

Our website is www.theceme.org.

For further information please contact the Director, Revd Dr Richard Turnbull, at:

The Centre for Enterprise, Markets and Ethics
First Floor, 31 Beaumont Street,
Oxford OX1 2NP

About the Author

Andrei E. Rogobete is Associate Director at the Centre for Enterprise, Markets and Ethics (CEME). His main area of research addresses the moral foundation of the adoption and impact of AI within the field of business ethics. He applies a distinctive theological perspective rooted in Judaeo-Christian teaching. Andrei is the author of several publications, including *Taxing Families Fairly* (co-authored with Prof. Philip Booth, published by Policy Exchange), *Ethics in Global Business* and *The UK Savings Crisis*. Andrei holds a CTS in Theology from Oxford University, an MSc in Business Management from University College London (UCL) and a BA (Hons) in Political Science and International Relations from Royal Holloway, University of London.

Contents

Introduction

Spike Jonze's 2002 science-fiction film *her* depicts an unusual romance between Theodore Twombly, the main protagonist, and Samantha, a superintelligent virtual assistant with artificial intelligence (AI). Theodore becomes so enthralled by Samantha's ability to learn from human interactions and events that the two form an emotionally charged relationship of intrigue, affection and vulnerability. For Theodore, Samantha is as sentient as any human being despite lacking a physical body. She praises Theodore for helping her discover the characteristically human-like desire of wanting.[1] Yet the tension between appearing sentient and being sentient thematically permeates much of the background of their unconventional interactions. Theodore often finds himself wrestling with the human-but-not-quite dimensional reality of the relationship. It all ends abruptly when Samantha switches off, leaving Theodore dismayed yet in some sense content with the entire experience.

Humanity's search for what resembles itself is not a novel phenomenon – the near obsession with creating or discovering objects that bear human attributes dates back to antiquity. Ancient Greece produced the mythological figure of Talos, a giant bronze automaton created by Hephaestus to protect Europa from foreign enemies.[2] Throughout history the elemental substrates that comprise intelligence – such as consciousness, creativity and wisdom – have generated points of contention and of intrigue.

In contemporary times, the advent of increasingly complex forms of AI raises a significant question: Are we at the cusp of a new technological era or is it another fanaticised case of AI hyperbole? What is becoming clear is that embedded AI is proliferating into a growing number of services and products, with profound implications for both public and private spheres. For businesses this will probably result in efficiency and productivity gains over the short to medium term. The longer term will most likely see growth in such areas as large language models and generative AI more broadly, leading to significant changes in use of and reliance on AI-driven automation for products, services and data analysis. Commonly established processes are expected to be transformed and streamlined. UBS Bank estimates that average annual growth in the AI sector will be around 20 per cent, affecting 50–75 million jobs. Some sectors that are traditionally slow adopters, such as retail and manufacturing, may be detrimentally impacted, while faster adopters, such as software, healthcare research, hi-tech engineering and select service companies, are more likely to experience favourable outcomes.[3]

The recent growth in the field of advanced AI systems is largely the culmination of progress in computer science and emergent algorithmic technologies over the last three decades. The spread of machine learning, deep learning and artificial neural networks has gradually transitioned from IT labs to even the most menial daily tasks. Human environments are experiencing this seismic shift driven by generative algorithmic computations. For the lay audience, what makes machine learning distinctive from previous forms of computer intelligence is that it brings the capacity both to compute with pre-existing data and actively learn from new data and generate information reliant on it.

The implications, though considerable, might not always be apparent, which has led to a plurality of views among experts. The computer scientist and cognitive psychologist Geoffrey Hinton says: 'I think people need to understand that deep learning is making a lot of things, behind-the-scenes, much better.'[4] A group of researchers from Harvard University point out that: 'algorithms don't always work smoothly. They don't always make ethical or accurate choices.'[5] We also see discontent among prominent figures on the future of AI more broadly. Elon Musk, the founder of Tesla and SpaceX, puts it bluntly: 'With artificial intelligence we are summoning the demon.'[6] Similarly the late Stephen Hawking in an interview with the BBC:

> The development of full artificial intelligence could spell the end of the human race … It would take off on its own, and re-design itself at an ever increasing rate. Humans, who are limited by slow biological evolution, couldn't compete, and would be superseded.[7]

Other computer scientists are less convinced concerning apocalyptic outcomes. On sentient AI, so Oriol Vinyals, a machine-learning scientist at Google DeepMind:

> any of the current models, although very useful and very good, I think [are] quite far from that [human-level sentience] … Looking at biological systems as opposed to these computer computational brains, there is such a difference in the level of complexity still, orders of magnitude of complexity … It just feels like it is not possible to achieve the same levels of complexity.[8]

What to make of these contrasting views?

It is certain that the advancement of increasingly complex forms of AI demands questions be asked about what it means to be truly human, as well as re-examination of the essence of human nature and humanity's place within creation. This raises profound ontological questions: What exactly does it mean to be human? What separates humans from the rest of creation? Are humans mere biochemical entities designed for solely computational tasks? What is the role and responsibility of humanity towards the created world and how should this shape humanity's interaction with it?

Judaeo-Christian teaching points to human beings as bearers of the *imago Dei*, the image and likeness of God. They are rooted in the transcendent, created for a higher purpose: a living relationship with the divine that surpasses the temporal, moving towards the metaphysical. An outright rejection of the metaphysical risks a more widespread fall into nihilism. Rooted in the Holy Trinity, which is ultimately defined by supreme interpersonal love, humanity is capable of experiencing the depths of such love, overflowing from the divine and extending into fellowship with the whole of humanity. This human–divine synergy allows and entails a process of continual growth and transformation, making it a distinct feature of what it means to be truly human.

From a Judaeo-Christian perspective therefore, it is false to presume that AI will ever have the capacity for spiritual connection with the transcendent God, or indeed with any other human being, regardless how complex or convincing it may appear. AI cannot transcend the material realm, thus moving or even aspiring towards the metaphysical. The limitation is the material world, bound to physics and matter. The adoption of AI also has no bearing on making humanity any more or less human. By definition, artificial intelligence cannot and will not equal the biochemical physical build of a human. This is not to deny that AI will bring substantial benefits to humanity – it already has and will continue to do so. Rather it is a matter of taking proper perspective, which then enables appropriate responses to AI.

The contention here is that the future relationship between AI and human beings will require a certain degree of discernment, understanding and wisdom. The first necessity is to ensure that AI is a net contributor to positive outcomes for human beings, society and the rest of creation. The second is to ensure new technologies are rooted in a set of fundamental ethical values. Many of these could be drawn from widely held beliefs, including human dignity, the respect and protection of life, freedom of speech, fairness in the treatment of others, non-maleficence and the rule of law. The implementation of moral values into the AI code should come through, and in some cases from, the programmers themselves. The third necessity is to promote a regulatory framework that looks at the foundational level and remains focused on the big picture. The dual challenge here is to safeguard the promotion of innovation while upholding a set of ethical standards. As will be argued, lessons from Judaeo-Christian teaching can furnish the debate with a strong moral framework.

The following study is divided into three chapters. Chapter 1 considers some of the milestones in AI history, highlighting the gradual advance from purely computational AI to AI with advanced features, such as machine learning, deep learning and pattern recognition. Chapter 2 presents two case studies. The first is an analysis of AI-led biometric identification, which showcases significant efficiency and security benefits brought to points of transit, such as national borders. The second considers the development of autonomous vehicles and how, though innovative, they are not yet deemed safe for widespread public use. Chapter 3 considers a Judaeo-Christian response to the adoption and use of AI. It emphasises the central role of humanity as image bearers of the divine, and the implications this brings for sentience as a uniquely human feature. (Terms such as 'artificial general intelligence', 'superintelligent AI' or 'general AI' will be used interchangeably.)

NOTES TO INTRODUCTION

1 IMDB Transcript Quotes, *her*, https://www.imdb.com/title/tt1798709/characters/nm0424060.
2 'Stanford Researcher examines Earliest Concepts of Artificial Intelligence, Robots in Ancient Myths', *Stanford News*, 28 February 2019, https://news.stanford.edu/2019/02/28/ancient-myths-reveal-early-fantasies-artificial-life/.
3 'AI's Coming of Age: The Beginning of True Autonomy', *UBS Bank*, https://www.ubs.com/microsites/artificial-intelligence/en/ai-coming-age.html.
4 Jose Fumo, 'Meet the Heroes of Deep Learning', *Towards Data Science*, 13 September 2017, https://towardsdatascience.com/meet-the-heroes-of-deep-learning-648c9083ef10.
5 Boris Babic, I. Glenn Cohen, Theodoros Evgeniou and Sara Gerke, 'When Machine Learning goes off the Rails: A Guide to Managing the Risks', *Harvard Business Review*, January–February 2021, https://hbr.org/2021/01/when-machine-learning-goes-off-the-rails.

6 'Tesla's Elon Musk: We're "Summoning the Demon" with Artificial Intelligence', *YouTube*, 24 November 2014, https://www.youtube.com/watch?v=Tzb_CSRO-0g.
7 'Stephen Hawking warns Artificial Intelligence could end Mankind', *BBC News*, 2 December 2014, https://www.bbc.co.uk/news/technology-30290540.
8 'Oriol Vinyals: Deep Learning and Artificial General Intelligence', Lex Fridman Podcast #306, *YouTube*, 26 July 2022, https://www.youtube.com/watch?v=aGBLRlLe7X8.

Chapter 1

The evolution of artificial intelligence

1.1 Defining the terms

Within academic and scientific spheres, the concepts surrounding artificial intelligence are often confined to an understanding of intelligence that is largely computational in nature and scope. In its most basic form, AI can be understood as an intentional action aimed at producing a certain outcome. Sometimes the input/output data is known, sometimes unknown – at least until the processing is complete. Yet the tendency remains to evaluate AI largely in relation to human intelligence, which goes beyond mere electronic processing and considers wider, sometimes unexpected consequences. Margaret Boden, Professor of Cognitive Science and Informatics at the University of Sussex, argues that:

> Artificial Intelligence seeks to make computers do the sorts of things that minds can do … Intelligence isn't a single dimension, but a richly structured space of diverse information-processing capacities. Accordingly, AI uses many different techniques, addressing many different tasks.[1]

However, viewing AI in relation to human intelligence presents significant challenges that much of the literature has struggled to explain or classify. For instance, if AI is to mimic human intelligence, how does this relate to conscience? Unlike the electronic calculator, human intelligence performs a simultaneous array of calculations alongside the given task. Again, some of these calculations are known, others are not. This raises challenges by making human intelligence less predictable, less quantifiable and consequently harder to emulate within traditional AI systems.

Since the turn of the century a distinction has been growing between narrow AI and general AI, or artificial general intelligence (AGI).[2] Narrow AI is the simple, single-task operation conducted by AI. It follows a concrete set of instructions and usually gives a predictable result. Examples include computer games, simple data analysis, certain algorithms – such as email filtering – and other applications encountered in day-to-day life yet often remaining hidden. Narrow AI 'displays a certain degree of intelligence in a particular field. It performs highly specialized tasks for humans, within that narrow field.'[3] More recently, narrow AI has evolved to encompass learning abilities, known as machine learning or deep learning. This involves not just executing a set task or goal but also learning from new data and adapting the execution based on input data – not a predetermined computation.

Artificial general intelligence, however, remains a hypothesis. It is a far more ambitious project that seeks to replicate the human brain in the entirety of its complexity, and potentially surpass it. The distinction between narrow AI and AGI is significant because the latter has intrinsic agency to act, a capacity to learn and an ability to perform actions based on the data it generates. Colloquially speaking, the machine has a mind of its own. Thinkers such as Yuval Noah Harari and Jaan Tallinn are proponents of this apocalyptic existential threat of an AI 'singularity' point in time, whereby the machine has improved itself to such an extent that it surpasses human intelligence and overpowers the human race.[4]

Yet there is a crucial shortcoming: unlike the human mind, AGI fundamentally lacks non-computational thinking – or conscience – and a deeper sense of self-awareness that goes beyond algorithmic learning. By default, therefore, AGI cannot possess any inherent gauge of morality except for general rules either preprogrammed or learned, which raises key existential questions: Is sentience something that can be programmed or does it originate from an external source? If, hypothetically speaking, consciousness can be programmed, would it be genuine consciousness (in a human sense)? Indeed, what is genuine consciousness? Would it even matter as long as there are positive and 'moral' outcomes? John Lennox, Professor of Mathematics at Oxford University, has this to say about genuine sentience:

> it is one thing to make a machine that can simulate, say, a human hand lifting an object; it is a completely different thing to make a machine that can simulate the thoughts of a human when he or she is lifting an object. It is much easier to do the first than the second, and if utility is all that is required, then the first is all that is necessary … For AI computer systems have no conscience, and so the morality of any decisions they make will reflect the morality of the computer programmers – and that is where the difficulties start.[5]

This lack of embedded morality or the ability for self-reflection against a moral code is indeed where the difficulties start for complex systems such as AGI. Could an AGI compose a poem about war? It most likely could – and in some sense current, generative AI already has. The Appendix offers two poems generated by ChatGPT, an artificial intelligence chatbot developed by OpenAI and launched in November 2022. The two poems are on the topics of consciousness and war and make for intriguing reading.

However, at what level would a superintelligent AI – or ChatGPT currently – engage with the topic of war? Indeed, can feelings and emotions about the gravity of war be manually implemented or, in the case of machine learning, can an AGI be encouraged to learn feelings of sorrow or loss? If it does express feelings of sorrow or loss, would these be anything more than computational? Would it even matter as long as humans are unable to distinguish between the two? Chapter 3 returns to some of these questions, discussing this interrelationship in greater depth.

Since the 1980s, AI's *utility* has moved from being its principal objective to one of many. Creativity and innovation produced by AI itself are now held in higher regard and seen as more enticing to programmers than utility alone. Professor Stuart J. Russell points out that historically there are two different approaches to AI.[6] One is viewing AI in 'fidelity to human performance', the other is a more abstract 'definition [of intelligence] called rationality – [or] loosely speaking, doing the "right thing"'.[7] The two therefore give rise respectively to artificial systems that aim to mimic humans in both behaviour and output, and those primarily rational in behaviour and output. These differing approaches lead to rather different pursuits of recreating intelligence, whether akin to human intelligence or not – demonstrating the breadth and depth of the issue.

Dr John McCarthy from Stanford University defines AI as:

> the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.[8]

He argues that there is an entire list of branches within the field of AI, many not yet clearly defined or even apprehended. Areas of enquiry include, inter alia, logical AI, pattern recognition, representation, inference, heuristics and genetic programming.[9]

Attempting to define AI brings its own set of challenges, not least because the continually evolving space results in an ever-growing number of known unknowns, as well as unknown unknowns, or future pathways of AI evolution that scientists have not yet thought about. It also brings further challenges because the distinction between narrow AI and AGI does not give sufficient weight to the sheer complexity and difficulty of the latter. To create something that can rival the human brain, not just in utility but also in a sense of conscience, situational awareness and the biochemical-emotional relationship, remains no more than an aspiration. For now, we can define artificial intelligence as the leveraging of computer power – in all forms – to computational tasks such as learning, problem-solving and data generation.

1.2 Milestones in AI history

The idea of creating intelligent machines dates back to antiquity. For millennia, humanity has been intrigued – in some cases obsessed – with the idea of creating objects that to a greater or lesser degree reflect their own image. The word 'automation' comes from the Greek *automatos*, meaning 'self-acting, moving or acting on its own'.[10] It was coined by the Greek-Egyptian mathematician and inventor Heron of Alexandria (*c.*AD 0–70), when he created the first set of automated temple doors. Driven by a steam-powered hydraulic mechanism, the doors would open when the fire to the temple altar was lit and close when it was extinguished – much to the astonishment of devotees.[11] In his writings and early sketches Heron named the invention 'Number 37', marking an important milestone because it represents – albeit this is difficult to prove conclusively – the earliest form of a human-engineered automated machine. Number 37 was the first mechanism with the ability to act on its own.

Other objects considered at the time to have some degree of intrinsic intelligence were principally objects of divine worship. Their capabilities were wholly or in part derived from their divine status. Belief in them required faith, not observation. The Introduction mentioned Thalos, a giant automaton of bronze built by Hephaestus, the Greek mythological God of invention and blacksmithing. Thalos was ordered by Zeus, king of the Gods, to patrol Crete three times a day to protect it from foreign invaders. At the core of his body Thalos held his mysterious life source, something the Greeks called 'ichor'. Mythology literature reveals how Thalos was killed by a sorcerer who pulled a bolt from his ankle, which caused the ichor fluid to flow out.[12] Yet Thalos was a myth, and his existence in the imagination of the ancient Greeks illustrates the powerful allure of machines that embody superior human features. For Greeks, Thalos represented what they lacked themselves: a large and powerful human-like machine that reaped the benefits of technological advancement and was blessed by the gods – in this case Zeus – to offer protection from enemies.

Leaving aside objects such as figurines or statues given divine attributes, the abacus is often regarded as the first intelligent machine from a computational standpoint. The earliest archaeological evidence of the abacus dates back to the third millennium BC, within the Babylonian empire and the Early Dynastic period in Egypt. Ettore Carruccio, a scholar of the Babylonian period, pointed to evidence that Babylonians 'may have used the abacus for the operations of addition and subtraction'.[13] Though simple by contemporary standards, the abacus was probably the first tool that enabled a visual representation of basic calculations, thus reducing the possibility for error and expanding understanding of elementary mathematics to a wider audience. Its impact was vast and to some extent the abacus can be viewed as a primitive form of personal computing device.

Moving briskly forward to the early 1800s, we arrive at the British mathematician, philosopher and inventor Charles Babbage, who proposed the Analytical Engine. Although Babbage himself never had the opportunity to complete the build, his design marked a milestone in the field of computer science – he was a pioneer in the development of intelligent machines. His Analytical Engine was a much-improved successor to his previous design, the Difference Engine, and it achieved what was later known as 'turning complete'. This meant that it incorporated the same basic logical structures that electronic computers would later possess. The Analytical Engine comprised of an arithmetic logic unit, 'control flow' in the form of conditional branching and loops, and integrated memory – making it the first general-purpose programmable computing engine.[14] Babbage became known as the father of computing for his remarkable skill in precision engineering, creativity of design and intellectual curiosity.[15] This paved the way for computer science to develop not just in its practical applications – which were themselves useful even in the early days – but as a recognised academic field of inquiry. Machines such as the Difference Engine were mostly used for applications that required tabulating polynomial functions. Despite Babbage's remarkable contributions to the field, he failed to secure sufficient funding to support his work and bring many of his projects to completion which, according to one source, 'left Babbage in his declining years a disappointed and embittered man'.[16]

Another key figure in the history of AI was Alan Turing. If Charles Babbage was the father of computing, Turing took that role for modern computer science.[17] Without going into detail on Turing's many accomplishments, it is worth mentioning some of his key contributions. 'The genius of Bletchley Park', Alan Turing is best known in popular culture for cracking the German Enigma code during the Second World War. It is thought that Churchill once credited him with making the single biggest contribution to the Allied victory against Nazi Germany. President Eisenhower noted in a letter that the work at Bletchley Park 'saved countless British and American lives and, in no small way, contributed to the speed with which the enemy was routed and eventually forced to surrender'.[18]

But concerning artificial intelligence, Turing produced two seminal papers that effectively established the field of computer science. What was novel about his thought was that by as early as 1937 he imagined a machine that could be instructed by multiple computations with consequentially varied outcomes. This in turn could be used to resolve a diverse array of computational problems. It was in some sense the earliest feasible description of a multipurpose and multitasking computer. At a time when machines were limited to singular preprogrammed computations, Turing's first paper – 'On Computable Numbers, with an Application to the Entscheidungsproblem' – was revolutionary.[19] Professor Joel Hamkins argues that:

> The paper is an incredible achievement. He accomplishes so much: he defines and explains the machines; he proves that there is a universal Turing machine … he argues that his machine concept captures our intuitive notion of computability; and he develops the theory of computable real numbers.[20]

In his mathematical descriptions Turing called it the Universal Machine, though it would later become more commonly known as the Turing Machine. What makes the machine special is that it could, in theory, solve any mathematical problem presented to it in symbolic form. This is largely because Turing developed it along similar lines to a human carrying out mathematical computational processes. Yet some computer scientists and mathematicians, such as Professor Martin Davis, argue that the true legacy of Turing's Universal Machine is the development of the 'the stored-program computer': placing the input data as well as the machine instructions on a uniform memory, something present today in all computer devices.[21] Paul Gray from *Time* magazine wrote that 'everyone who taps at a keyboard ... is working on an incarnation of a Turing machine.'[22]

The second Turing paper to touch on here was 'Computing Machinery and Intelligence', published in 1950.[23] It postulates the question of machine intelligence – Can machines think? – and attempts to demonstrate intelligence via an imitation game: Can AI become indistinguishable from human intelligence and can it fool a human into thinking it is human itself? This would commonly become known as the Turing test.

Illustrated in Figure 1, the Turing test involves three actors: two guests and an interrogator. One of the guests is the AI (person A) while the other is a human (person B). The AI has to trick the interrogator (person C, also human) into thinking it is human. The 'conversation' has a duration of five minutes and if the AI successfully convinces the interrogator it is human over 30 per cent of the time, the test has been passed.

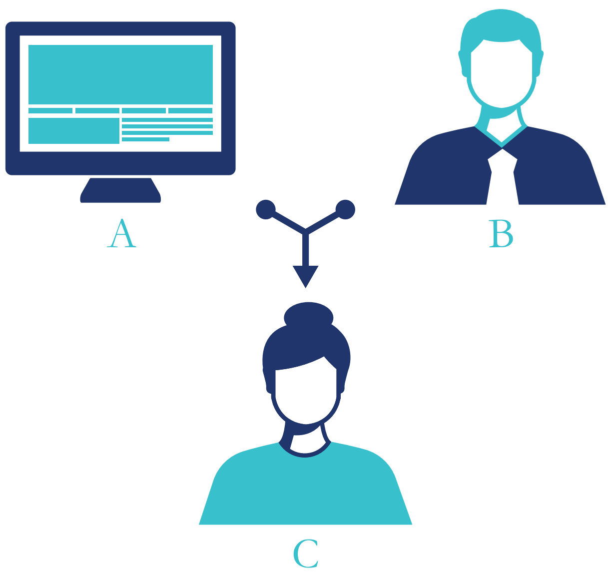


FIGURE 1: THE TURING TEST

The interview consists of a series of questions and answers whereby the interrogator probes A and B individually. One of the major attractions of the test is its simplicity and flexibility: it allows the interrogator to pose a wide variety of questions and intellectual – or emotional – challenges. Turing himself wrote that: 'The question and answer method seems to be suitable for introducing almost any one of the fields of human endeavour that we wish to include.'[24]

The test consequently had a significant impact not just on the development of artificial intelligence itself but on how we measure AI in relation to human intelligence. It sparked the debate on our capacity – or incapacity – to quantify and evaluate intelligence, itself an abstract, relative term. The test therefore represents more of a conceptual exercise than a scientific measuring tool. Professor Daniel Dennett has said that Alan Turing intended it as 'a conversation stopper, as a thought experiment that should convince people that any computer that would pass this test – of course it would be intelligent'.[25]

One important challenge of the Turing test is its reliance on humans as judges in classifying the subject in question. This has caused some unreliable results because humans can be poor at recognising other humans through text alone. They often fall prey to what is known as the Eliza effect; that is, our subconscious tendency to view computer intelligence as analogous to human intelligence – a form of anthropomorphisation (in this case, of computer systems).

A good example of the anthropomorphic Eliza effect is given by the computer scientist Douglas Hofstadter, and involves a cash machine that displays the words 'Thank You' at the end of a transaction. The human may consciously or subconsciously think that it is expressing gratitude when all the machine is doing is showing a preprogrammed automated response.[26] The Eliza effect can thus lead to some unintelligent machines passing the test. For instance, delaying the response or remaining silent can be misinterpreted as thought processing. There have even been cases in which humans were misidentified as machines – known as the Confederate effect.[27] Here Kevin Warwick and Huma Shah have conducted research and found that, rather counterintuitively, under testing conditions humans attempting to 'act human' actually appear less 'human-like' and increasingly prone to giving computer-type responses.[28]

1.3 AI in the information age (1950s onwards)

To talk about the history of AI in the information age – or technological era – is, in effect, to discuss the history of computer science. The term 'artificial intelligence' was only coined during the 1956 Dartmouth College conference in Hanover, New Hampshire.[29] Following Turing's seminal work, the 1950s and 60s marked a period of great enthusiasm for machine intelligence. Turing himself believed that human-level artificial intelligence would be achieved within 50 years.[30] The MIT cognitive scientist Marvin Minsky, and others who attended the Dartmouth conference, were extremely optimistic about AI's future: within 'a generation ... the problem of creating "artificial intelligence" will substantially be solved'.[31] Scientists focused on challenges – mathematical or in the form of logic puzzles, games, pattern interpretation – that mimicked or were considered indicative of human intelligence, and attempted to meet them with AI.

Curiously, the personal computer (PC), when first invented, was not recognised as a milestone in the development of AI. In 1974 a small electronics company called MITS created the Altair 8800, the first PC to earn widespread popularity among the early tech community. Ed Roberts, the founder of MITS, claimed that as many as 40,000 units were sold, although the precise number remains uncertain.[32]

What the Altair 8800 did was spark a revolution in demand for personal computing. This came at a time when many doubts were cast over the purpose and usefulness of PCs within the household. 'Computerphobia', as it was known in the 1970s and 80s, became a real issue: 'computers won't make toast or vacuum a carpet', wrote the journalist Charles Rubin in 1983.[33] Early research suggested that fuelling this phobia were fears, among others, of damaging the computer, looking foolish, losing electricity and lacking control over the machine.[34] There was also a period in the early stages when large swathes of the public, as well as some experts, struggled to find much use for the PC. The computer scientist Ken Olsen famously said in 1977: 'There is no reason for any individual to have a computer in his home' (although Olsen was speaking in reference to computers being used for home automation).[35] Adoption of the PC and, by extension, the development of AI, therefore proved by no means straightforward, and slow progress as well as lack of funding led to the period 1974–80 becoming known as the first 'AI winter'.[36]

Despite these hurdles, the early 1980s were marked by a rebound in the development of new hardware as well as more intuitive, user-friendly software. The Xerox 8010 Star, released in 1981, was the first commercial personal computer to incorporate many of the core technologies that PCs feature today, including a graphical user interface, an operating system with icons and folders, a two-button mouse, Ethernet connectivity and even email.[37]

The Star's graphical interface became known as What You See Is What You Get (WYSIWYG), where users were encouraged to take actions directly on screen via the mouse and keyboard. For the first time, visual elements such as icons, taskbars, folders and documents represented interactive objects. WYSIWYG was revolutionary not just because it implemented what was then considered a seamless interface but because it offered the new possibility for those without prior knowledge of coding to use and interact with the device. Many of the functions and abilities of the computer became self-taught through simple trial and error. The intuitive design language encouraged users to think of the icons on the screen the same way they would their physical counterparts; that is, folders, trash, pictures, games and so on.[38] At the time, a young Steve Jobs was offered an opportunity to visit Xerox's Palo Alto Research Centre. Recalling the trip, Jobs said Xerox:

> showed me three things, but I was so blinded by the first one that I didn't really see the other two ... The graphical user interface, I thought, was the best thing I'd seen in my life ... Within 10 minutes it was obvious to me that all computers would work like this, someday. It was obvious.[39]

Steve Jobs's Apple Macintosh went on to sell a record 4.5 million computers in 1995 alone.[40]

1.4 Current trends in AI

The distinction between narrow AI and AGI was considered above, noting how narrow AI represents simple, often single-task computation that follows a concrete set of instructions and usually gives a predictable result. Much of the discussion so far has been of narrow AI, but now turns to an expansion of it that first emerged in the 1980s known as machine learning (ML).

Dr Issam Naqa defines ML as 'an evolving branch of computational algorithms that are designed to emulate human intelligence by learning from the surrounding environment'.[41] IBM sees it as 'a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy'.[42] Generally accepted definitions of ML bear certain similarities. Therefore a working definition that encompasses the fundamentals might say that ML is a branch within AI where computer systems have the ability to learn from new data and perform specific tasks with it. This data can be acquired from external sources or generated by the program itself, as would be the case in generative AI, for example.

One notable advancement in the field of ML learning came in 1981 when Professor Gerald DeJong proposed explanation-based learning (EBL). Although DeJong himself admits that 'there is yet no satisfactory answer to the question' of what exactly EBL is, we can broadly see it as a strategy of ML that learns from examples and has the capacity to draw generalisations or form concepts.[43] Stuart Russell defines explanation-based learning as 'a method for extracting general rules from individual observations'.[44] For instance, EBL has the capacity to create a program that learns to play chess by taking training examples and determining what are the relevant features in order to form a generalisation. EBL is usually based on four inputs: a hypothesis, a domain theory, training examples and the operationality criteria (determining which features are efficiently recognisable).[45] It is important because it represents one of the earliest forms of machine learning that attempts to imitate the human process of learning; that is, developing conclusions through the gathering of information and utilisation of prior examples.

Deep learning is a further evolution of machine learning that has dominated discourse in more recent times. If machine learning represents the ability to learn without explicit programming, deep learning takes matters a step further and utilises the gathered data to extract conclusions via neural networks. In this sense, it represents a more complex subset method of machine learning. Yet the term itself encompasses a broader variety of techniques, which are centred around algebraic circuits – 'deep' refers to the fact that in many cases the circuits have multiple layers (akin to a human neural network), making the computational path from input to output in many cases quite extensive.[46]

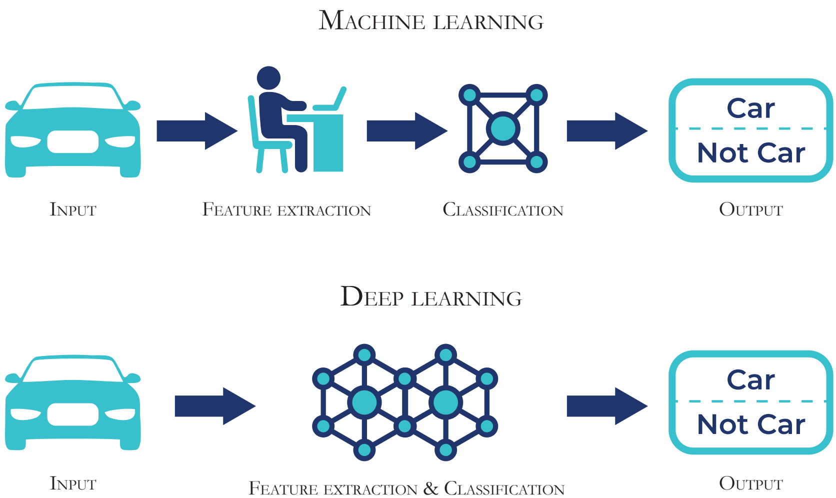


Figure 2: Differences between machine learning and deep learning

Figure 2 illustrates a basic example of how deep learning works in practice. Suppose an AI can differentiate between what is and is not a car. Under machine learning the features of what make up a car are manually programmed into the algorithm – this is known as 'feature extraction'. Under deep learning the artificial neural network picks up the distinct array of features itself without the need of manual input:

it does not require predetermined 'labelled data' to draw conclusions. However, this increased degree of autonomy demands a greater volume of data. Thus the AI is able to learn the characteristic features of a car through the automatic processing of an extensive number of examples.

From a functional standpoint, machine learning in general and deep learning in particular carry fundamental implications for how we interact with our work, social activities, families and even ourselves. Algorithmic calculations dictate what we see and often what we do not see in the virtual sphere. Yet machine learning is all the more pertinent precisely because new ML technologies such as generative AI are becoming increasingly embedded features of contemporary life.

This chapter has briefly sketched the history and some of the key milestones in the development of AI, as a basis for considering the complex implications of its utilisation and adoption. Chapter 2 takes a more practical approach by considering two case studies in AI.

Notes to Chapter 1

1 Margaret Boden, *Artificial Intelligence: A Very Brief Introduction*, Oxford: Oxford University Press, 2018, p. 1.

2 'Narrow AI vs. General AI – What's Next for the Future of Tech?', *Analytics Insight*, 5 February 2022, https://www.analyticsinsight.net/narrow-ai-vs-general-ai-whats-next-for-the-future-of-tech/.

3 'Narrow AI vs. General AI', *Analytics Insight*.

4 Benjamin Charles Germain Lee, 'The Singularity Prophets', *Current Affairs*, 23 July 2020, https://www.currentaffairs.org/2020/07/the-singularity-prophets.

5 John C. Lennox, *2084: Artificial Intelligence and the Future of Humanity*, Grand Rapids, MI: Zondervan, 2020, pp. 17, 145.

6 Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, Global edition, Harlow: Pearson, 2021, pp. 19–20.

7 Russell and Norvig, *Artificial Intelligence*, pp. 19–20.

8 Richard S. Sutton, 'John McCarthy's Definition of Intelligence', *Journal of Artificial General Intelligence* 11:2 (2020), pp. 66–7.
9 Sutton, 'John McCarthy's Definition of Intelligence', pp. 66–7.
10 Online Etymology Dictionary, 'automatic', 2023, https://www.etymonline.com/word/automatic.
11 'Heron of Alexandria – Automated Temple Doors', *Artefacts – Scientific Illustration and Archaeological Reconstruction*, 2021, http://www.artefacts-berlin.de/portfolio-item/heron-of-alexandria-automated-temple-doors/.
12 'Stanford Researcher examines Earliest Concepts of Artificial Intelligence, Robots in Ancient Myths', *Stanford News*, 28 February 2019, https://news.stanford.edu/2019/02/28/ancient-myths-reveal-early-fantasies-artificial-life/.
13 Ettore Carruccio, *Mathematics and Logic in History and in Contemporary Thought*, trans. Isabel Quigly, Piscataway, NJ: Aldine Transaction, 2006, p. 14.
14 'The Engines', *Computer History Museum*, 15 January 2023, https://www.computerhistory.org/babbage/engines.
15 University of Minnesota, Charles Babbage Institute, 'Computing, Information and Culture: Who was Charles Babbage?', *College of Science and Engineering and University Libraries*, https://cse.umn.edu/cbi/who-was-charles-babbage.
16 University of Minnesota, Charles Babbage Institute, 'Computing, Information and Culture'.
17 'Alan Turing: The Father of Modern Computer Science', *The Twickenham Museum*, 2023 http://www.twickenham-museum.org.uk/detail.php?aid=66&ctid=1&cid=13.
18 Patrick Sawer, 'Letter reveals Bletchley Park Code Breakers secretly thanked by General Eisenhower for "Priceless" Work', *Eisenhower Foundation*, 15 March 2016, https://www.eisenhowerfoundation.net/sites/default/files/2020-05/File%206%2C%20Article%201.pdf.
19 Alan Turing, 'On Computable Numbers, with an Application to the Entscheidungsproblem', *Proceedings of the London Mathematical Society* Series 2, 42:1 (1937), pp. 230–65.
20 Joel David Hamkins, 'Alan Turing, On Computable Numbers', *Joel David Hamkins*, 29 September 2018, http://jdh.hamkins.org/alan-turing-on-computable-numbers/.
21 Martin Davis, *The Universal Computer: The Road from Leibniz to Turing*, New York: Norton, 2000, p. 121.

22 Paul Gray, 'Computer Scientist – Alan Turing: While addressing a Problem in the Arcane Field of Mathematical Logic, he imagined a Machine that could mimic Human Reasoning. Sound familiar?' *Time Magazine Archives*, 29 March 1999, https://content.time.com/time/subscriber/article/0,33009,990624-3,00.html.

23 Alan Turing, 'Computing Machinery and Intelligence', *Mind* 59:236 (1950), pp. 433–60.

24 Turing, 'Computing Machinery and Intelligence'.

25 'Turing Test – Daniel Dennett', *YouTube*, 22 May 2022, https://www.youtube.com/watch?v=e8vZy8a9lSc.

26 Douglas R. Hofstadter and the Fluid Analysis Research Group, 'Preface 4: The Ineradicable Eliza Effect and its Dangers, Epilogue', in *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*, New York: Basic Books, 1996, p. 157.

27 Huma Shah and Odette Henry, 'The Confederate Effect in Human–Machine Textual Interaction', conference paper, 2005.

28 Kevin Warwick and Huma Shah, 'Human Misidentification in Turing Tests', *Journal of Experimental and Theoretical Artificial Intelligence* 27:2 (2015), pp. 123–35.

29 Jørgen Veisdal, 'The Birthplace of AI: The 1956 Dartmouth Workshop', *Medium*, 12 September 2019, https://www.cantorsparadise.com/the-birthplace-of-ai-9ab7d4e5fb00.

30 H. P. Newquist, *The Brain Makers: The History of Artificial Intelligence – Genius, Ego, and Greed in the Quest for Machines that Think*, New York: The Relayer Group, 2020, p. 51.

31 Daniel Crevier, *AI: The Tumultuous Search for Artificial Intelligence*, New York: Basic Books, 1993.

32 Adwater & Stir, 'The First?', 15 February 2023, https://adwaterandstir.com/altair-history/.

33 David Neil Crocker, 'Excerpts from "Some People *should* be afraid of Computers", but Modernized', *David Neil Crocker*, 26 March 2019, http://crockedile.com/2019/03/26/textual-poaching-computerphobia/.

34 Adrienne LaFrance, 'When People Feared Computers', *The Atlantic*, 30 March 2015, https://www.theatlantic.com/technology/archive/2015/03/when-people-feared-computers/388919/.

35 Alex Bracetti, 'The Spam Crisis Solved? The 25 Craziest Things ever said by Tech CEOs', *Complex*, 14 January 2013, https://www.complex.com/pop-culture/2013/01/the-25-craziest-things-ever-said-by-tech-ceos.

36 Jim Howe, 'Artificial Intelligence at Edinburgh University: A Perspective', 2007, https://www.inf.ed.ac.uk/about/AIhistory.html.

37 'Xerox Star 8010 Information System, 1981', *The Interface Experience: 40 Years of Personal Computing* (Bard Graduate Center), 15 January 2023, https://interface-experience.org/objects/xerox-star-8010-information-system/.
38 'Xerox Star 8010 Information System, 1981'.
39 'How Steve Jobs got the Ideas of GUI from XEROX', *YouTube*, 4 January 2014, https://www.youtube.com/watch?v=J33pVRdxWbw.
40 Jeremy Reimer, 'Total Share: 30 Years of Personal Computer Market Share Figures – It's been a Long, Strange Trip for the Personal Computer over 30 Years', *Arstechnica*, 15 December 2005, https://arstechnica.com/features/2005/12/total-share/7/.
41 Issam El Naqa, Ruijiang Li and Martin J. Murphy, 2015, 'What is Machine Learning?', in El Naqa, Li and Murphy (eds), *Machine Learning in Radiation Oncology: Theory and Applications*, Cham: Springer, 2021, https://link.springer.com/chapter/10.1007/978-3-319-18305-3_1.
42 'Machine Learning: What is Machine Learning?', *IBM Cloud Learn Hub*, 11 January 2023, https://www.ibm.com/cloud/learn/machine-learning.
43 Howard E. Shrobe, *Exploring Artificial Intelligence: Survey Talks from the National Conferences on Artificial Intelligence*, Amsterdam: Elsevier Science, 2014.
44 Russell and Norvig, *Artificial Intelligence*, p. 750.
45 Richard M. Keller, 'Defining Operationality for Explanation-Based Learning', *Artificial Intelligence* 35:2 (1988), pp. 227–41.
46 Russell and Norvig, *Artificial Intelligence*, p. 801.

Chapter 2

Case studies: the opportunities and challenges of AI

Case studies are a useful way to evaluate the impact of AI on day-to-day life. The first one below considers the use of biometric recognition in strengthening cross-border security; the second the use of AI in autonomous driving vehicles. The aim is further understanding of where AI features, such as machine learning, might be beneficial and where detrimental.

2.1 Case study A: AI in the use of biometric security

2.1.1 Context and background

A prerequisite of effective governance is an awareness of who is being governed. Historically, the fingerprint represented the *de jure* form of authentication used by central governments and local authorities. Some of the first records of fingerprinting date back to the late nineteenth century, when William Herschel, a British colonial administrator in India, used basic ink to create impressions of the hands and fingers of awardees of civil contracts and administrative duties.[1] The motivation was essentially prevention of fraud and misidentification. Later, photographic identification, coupled with fingerprinting, increasingly dominated and advanced forensic criminology throughout the twentieth century.

A major breakthrough came in the 1980s, when IBM developed the first optical scanners, which could store scanned fingerprints into a digital database. This new technology extended to barcode scanning, which increasingly enabled categorisation of products and items throughout the late 1970s and 80s.[2] It facilitated the rapid transit and security screening of both goods and people. Digital databases ballooned in size and with them so did the stored information about individuals and products. Yet the more important aspect was speed of accessing data rather than capacity for storing it.

Criminals' fingerprints could now be screened and the results presented in real time. The implications for border security were, unsurprisingly, vast. By the mid-2000s, biometric passports – also known as epassports – had become commonplace (Malaysia was first to issue them, in 1998). By 2008, around 60 countries had biometric passports; by 2019, over 150.[3] Widespread adoption attests to the popularity and perceived benefits of biometric forms of identification compared to traditional documents.

2.1.2 Key issues

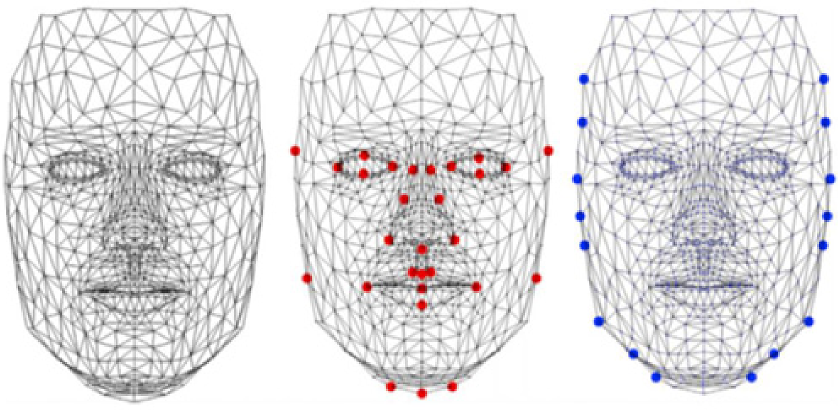
First to note is the role of national – or supranational – borders, namely to allow free passage of legal goods and people and prevent illegal activity. The UK Border Force defines its role as 'facilitating the legitimate movement of individuals and goods, whilst preventing those that would cause harm from entering the UK'.[4]

The objective of border control therefore fundamentally relies on the distinction between legal and illegal. The problem is that without technology that enables rapid identification, triaging between legal and illegal becomes a laboriously slow and ineffective process. This is particularly acute at high-volume borders, such as airports, seaports or major land border crossings. Guaranteeing safety as well as speed of passage remains a pressing, often conflicting issue. Automation in this case brings multiple benefits to those both seeking and overseeing transit. Biometric travel documents have also led to an increase in the number of automated border control systems, which allow passengers to pass through security merely by scanning their passport and/or face.

2.1.3 Action taken

The first British biometric passport was issued on 6 March 2006.[5] It featured a new design, with an array of supplementary tech-driven security features. A key distinguishing element is an embedded microprocessor chip that includes the holder's biometric data. The chip has the ability to store facial recognition, fingerprint recognition, iris recognition or any combination of the three. Underlying this technology is a system known as radio frequency identification, which enables the stored data to be encrypted and only read via the embedded radio tag itself. This means that no personal information can be retrieved and copied out of the chip.[6]

One advantage is that digitally stored data can be checked with more accuracy than is possible via visual inspection. Biometric facial data often includes measurements, such as distance between eyes, unique skin marks, size and shape of face. Figure 3 illustrates a facial-recognition mapping model.[7] This approach creates a network of geometric datapoints that can be identified on the surface of the face.



(**a**) Creation of the 3D face of a person

(**b**) Results of the detection of 29 landmarks of a face using the active shape model

(**c**) Results of the detection of 26 landmarks of a face

FIGURE 3: FACIAL RECOGNITION MAPPING

Facial recognition then uses key-points-based techniques to interpret and establish the geometric characteristics of the facial surface.[8] Again, these can include measurements surrounding the inner features of the face (such as nose, lips, eyes), as well as outer features (such as jawline, cheekbone contour, head shape). The facial-recognition process usually involves two steps at which machine learning is involved. The first is key-point detection, where the AI detects the defining physical features of the surface and assigns key points.[9] The second is feature extraction, where data is generated by mapping the key points. The final result is a full-scale digital model of a person's face, with over 80 key nodal points that can be accurately utilised to confirm identity.

2.1.4 Results and the impact of AI

The move to biometric passports marked a significant turn from traditional travel documents that largely relied on distinctive physical indicators for security and identification. The most obvious result was an increase in the level of border security. Facial recognition benefits from precise measurements as well as the ability to adapt to physical changes, including biological markers such as facial hair, aging or even skin tone. The benefit of having an extensive map of data points is that even if a certain percentage are altered or obstructed (such as by glasses), the remaining visible part of the facial map can be used to authenticate identity. At a micro level, the human face is so unique that the chance of misidentification is estimated to be around 1 in 1,000,000, compared to 1 in 50,000 for fingerprint recognition.[10] Peter Schmallegger of NXP Semiconductors argues that: 'The use of ePassports makes it harder to present doctored or illegally issued travel documents as the real thing, and helps reduce black-market trade in travel documents.'[11] The data held on the inside chip, compared to embedded security features on traditional passports, makes epassports difficult to duplicate.

Despite officers' training, visual inspections have their limitations – there will always be cases where the similarities are dominant and can lead to misidentification, such as with twins. Dominik Malčík and Martin Drahanský's extensive research in this regard found that:

> If the facial photo is treated from a biometric point of view (not just as a picture of a person) – the face contains information that is invariant in time and can be measured … for example, the distance between eyes, position of chin, position of nose, and so forth. These factors can affect the recognition process by providing additional information to the officer.[12]

Without the availability of biometric data, officers must rely solely on their own training and intuition.

Moreover, the increase in security has resulted in border-control forces adopting and implementing biometric technologies throughout their operations, with overwhelming support from their chief executives. Accenture conducted a survey of 91 border-agency leaders from Australia, Finland, France, Germany, Japan, Norway, Singapore, the UK and the USA, and found 92 per cent willing to adopt next-generation biometric technologies. Of those surveyed, 68 per cent believed these new technologies can help reduce risk and improve border security, and 54 per cent that they will enhance customer-service delivery.[13]

Another, perhaps more consequential result of biometric passports has been an increase in speed of passage via the adoption of new forms of border automation such as e-gates, or 'automated border control systems'. E-gates are automated border-control barriers that utilise the data stored in the biometric passport microchip, along with the facial scan or fingerprint taken at point of entry. The system electronically verifies the passport holder's identity and grants – or denies – passage. If there are issues in the biometric verification process, most scenarios result in the gates remaining closed and an immigration officer taking over. It is important to note that border officers permanently oversee the verification process, usually from a remote desk in proximity to the gates. This enables full control should issues arise.

Biometric identification is a good example of how artificial intelligence and the human component can work harmoniously together. The cohesion of AI, in the form of biometric technology, with the element of human oversight, reinforces speed and security. The two are mutually inclusive, resulting in positive outcomes for both travellers and border staff.

In this sense, e-gates bring additional benefits by reducing the administrative burden on border forces and allowing them to focus on higher-risk individuals. In the future, predicts Michael Petrov of Vision-Box, which has installed over 800 e-gates in more than 60 airports across Europe:

> the trend will shift toward the integration of automatic and manual immigration checks. Initial traveler data … would be verified at various stages of the trip. When travellers are securely known throughout their entire journey, the border control mission will be fully realized.[14]

2.1.5 Potential risks and drawbacks

There are arguably two main issues associated with the use of biometric technology in places of transit: system malfunction and privacy. Both are resolvable.

Statistically speaking, the possibility of a systems malfunction is exceptionally low. The latest biometric-recognition technologies have a demonstrated accuracy of 99 per cent, and benefit from an inbuilt error-detection system that instantly alerts overseeing officers to irregularities.[15] In addition, the likelihood of malfunction, attempted illegal passage and officer error all occurring at the same time remains infinitesimally small. This translates to a system with very robust safeguards in place.

With regard to privacy, e-gates collect personal information that usually includes name, gender, date of birth, passport number, travel plan, passport photograph and facial biometric templates. While some might perceive a potential privacy risk, the reality is that, under current legislation, passengers already give up personal details if they wish to travel, whether or not via biometrics. In addition, the data is only handled by authorities sanctioned to do so. It is, however, important to ensure that data is securely stored and not vulnerable to fraudulent attacks or leaks. One e-gate manufacturer pointed out that: 'We will only use or disclose the personal information for the purpose for which it was collected or as otherwise required or authorised by law.'[16] Privacy in this case is therefore rather a non-issue: passengers always forfeit information if they wish to travel.

2.2 Case study B: the use of AI in autonomous driving

2.2.1 Context and background

In 1939, the American industrial designer Norman Bel Geddes created Futurama, an iconic diorama that depicted what urban transportation might be like in the 1960s. It was no small affair: using scale models, it featured a large metropolis with deep suburbs, 50,000 vehicles, over 500,000 buildings, a million trees, fast-moving highways, sky-high buildings and a 140-ton conveyer system where guests could sit and admire the spectacle.[17] First presented at the New York World Fair for General Motors, Futurama became the exhibition's 'number-one hit',[18] with over 30,000 visitors, fascinated by the possibilities and implications of technological automation.[19]

Inside Futurama, Bel Geddes also presented the first concept model of an autonomous vehicle guided by radio-controlled electromagnetic fields, created by placing metal spikes alongside the roads and highways. General Motors took this concept and made it a reality in 1958 by creating an autonomous vehicle guided entirely by electric pick-up coils. The flow of current could be altered to control the vehicle and move it left or right.[20] It proved unscalable and ultimately infeasible, yet sparked the imagination of both car engineers and the general public as to what advanced automation might bring. There was a common cultural perception by the mid-century that the age of technology was in full swing. In reality, since the 1950s much of the progress has been limited to largely functional vehicle improvements, each decade seeing the target of fully automated vehicles pushed further back as new complexities arise.

2.2.2 Key issues

Since the late 2000s, two corporate giants have been at the forefront of autonomous-driving innovations: Alphabet (Google's parent company) and Tesla, though others, including GM and Honda, conduct smaller-scale R&D. Thus far all efforts have failed despite the billions invested.[21] Elon Musk admitted in an interview that the problem of self-driving is:

> way harder than I thought … what it comes down to at the end of the day is, to solve self-driving … you basically need to recreate what humans do to drive … which is to recreate optical sensors, eyes, and biological neural nets – that is how the entire road system is designed to work.[22]

In similar fashion, Anthony Levandowski, the co-founder of Google's self-driving car project, Waymo, has said that one would be 'hard-pressed to find another industry that's invested so many dollars in R&D and that has delivered so little'.[23]

At present, most vehicles offer some semi-autonomous features more aimed at driver convenience and general safety than autonomous driving per se. Examples include cruise control (more recently radar- and GPS-guided), crash mitigation systems (such as obstacle detection, airbag deployment and assisted braking), as well as other optional improvements (such as auto-dipping headlights, automated windscreen wipers and even a suspension system that adapts its firmness based on real-time GPS scanning of the route ahead, first implemented by Mercedes-Benz in 2013).[24] It is expected that such technologies will continue to improve and gradually feature across a wider range of vehicles.

So with a fully-fledged, road-legal self-driving car yet to be achieved, how far has autonomous driving technology come?

2.2.3 Action taken

The Society of Automotive Engineers currently divides levels of automation into the following categories:

> Level 0 – no automation
>
> Level 1 – driver assistance (hands on/shared control)
>
> Level 2 – partial driving automation (hands can be taken off the steering wheel)
>
> Level 3 – conditional driving automation (does not require continuous eyesight attention)
>
> Level 4 – high driving automation (does not require continuous mental attention)
>
> Level 5 – full driving automation (the steering wheel becomes optional)[25]

The most sophisticated road-legal cars today have the technological capacity to reach Level 3. However, safety legislation only permits this for limited periods (usually seconds), after which the driver is required to interact with the steering wheel and demonstrate control of the vehicle.

Some of the main technologies currently used in autonomous driving include audio/visual cameras, GPS, thermographic cameras, radar and lidar (light detection and ranging).[26] Figure 4 illustrates where most of the sensors would usually be placed on a vehicle.

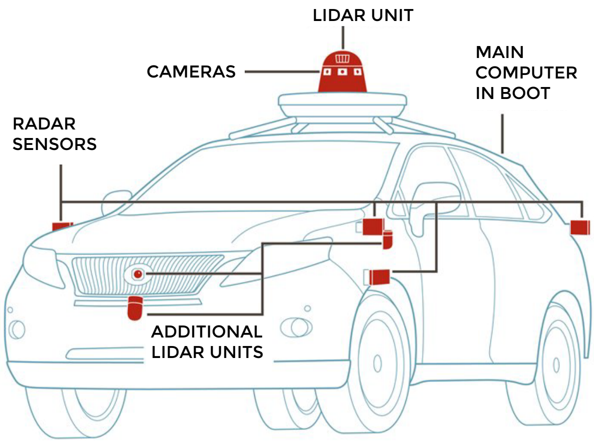


FIGURE 4: POSITION OF SENSORS ON A TYPICAL VEHICLE

Similar to pattern-recognition technology used in biometric identification, the depth- and object-perception sensors on the car are used to recreate a 3D digital model of the surrounding physical world. Not only does the 3D model have to be accurate, it needs to overcome the additional challenge of continuously adapting in real time.

Figure 5 illustrates what the digital map of the surrounding environment would look like. The challenge is for the systems to identify, interpret and act according to the patterns of objects and people presented while in motion. The AI must accurately label and categorise any familiar forms or shapes, which usually involves pedestrians, road signs, traffic lights, bus lanes, cycle lanes, trucks and other vehicles.

Figure 5: Lidar mapping of surrounding environment

In technical terms, the system used by the sensors is known as Bayesian simultaneous localisation and mapping (SLAM), though it is important to note that not all self-driving vehicles use SLAM. Algorithms in SLAM merge data inputs from the sensors placed throughout the car, which allows generation of a digital map similar to that shown in Figure 5.[27] Some autonomous-vehicle programs use a type of SLAM called DATMO (detection and tracking of other moving objects), which brings the additional advantage of identifying and predicting the position of an object once it has been defined (such as a cyclist, pushchair – or horse).[28]

2.2.4 Results and associated risks

Results and progress on autonomous driving thus far present a mixed picture. On the one hand, in June 2022, Cruise, a subsidiary of GM, received a Driverless Deployment Permit from the State of California – the first private company given permission to offer and charge for driverless rides within a major US city (San Francisco).[29] On the other, while the technology currently has the capability to perform basic driving- and crash-mitigation manoeuvres, it is not yet fit for mass adoption. There are three interconnected areas in which the main outstanding issues arise.

The first concerns when the navigation system encounters the unforeseen, such as a deer jumping in front of the car on a country road, or another vehicle rapidly changing lane without signalling. Information misinterpretation, as well as acute computational load on the system, can result in a temporary malfunction, placing the vehicle at risk of collision.

In March 2018, the death of Elaine Herzberg in Tempe, Arizona, marked the first fatal crash caused by a fully autonomous test vehicle on public roads.[30] Mrs Herzberg was pushing a bicycle across a four-lane carriageway when an Uber test car struck her via lateral collision at an estimated 45 mph (72 km/h).[31] While it became clear that the crossing itself was treacherous, the vehicle's CAS (collision avoidance systems) failed to activate. The AI's malfunction pointed to issues in the hardware (the sensors) as well as the software (the programming), which accordingly failed to engage. Jim McPherson, a California attorney and self-driving car expert, pointed out that: 'Lidar or Radar [systems], each has their shortcomings depending on light, reflectivity … Dark clothing and metal appear to one better than the other.'[32]

The second area concerns when the pattern-recognition AI wrongly identifies or interprets an object on the road (such as a red light or a mobility scooter). This usually occurs when the AI's recognition algorithms and/or the assortment of cameras and sensors see the data but either malfunction or mislabel it. There is evidence of an incident in 2016, when another Uber autonomous vehicle failed to recognise the correct colour at a traffic light and crossed on red.[33] One investigation suggests that the lasers and cameras used by driverless cars can misinterpret red traffic lights as green up to 30 per cent of the time.[34] Clearly, misinterpreting one colour for another while driving has potentially severe consequences.

The third area concerns when the AI fails to understand the unwritten rules of driving. A major component of driving relies on distinctly human anticipation and intuition. Making eye contact with other drivers, small hand gestures to pedestrians, giving way – or warning of hazards – by flashing headlights all contribute to a driving ecosystem heavily reliant on human perception and instinct. Current AI technology is poor in this area.

There are also issues surrounding the dynamics of object occlusion and object permanence. Human drivers rely on memory and quite intricate spatial awareness. For example, a driver encountering a school bus pulling over may only temporarily see the children disembarking but her short-term memory and intuition instinctively make her slow down in anticipation that some children may suddenly jump from behind the bus to cross the road. The object (in this case children) temporarily appears and then disappears from the driver's field of view, yet she remains aware that temporary disappearance does not imply absence – the object is merely obscured (in this case by the bus). An AI in such scenarios may be fooled into thinking that since the object is no longer in sight, it is no longer present.

Creating AI driving systems that can reliably mitigate the dangers of such intricate situations is work in progress. Google has invested heavily in R&D here with their 'behaviour prediction' technology, which 'can predict the behaviour of an object on the road based on its classification by inferring data from the training models constructed using millions of miles of driving experience'.[35] While promising, the software is not yet available to the public, and as even Google admit: 'This is an incredibly difficult problem that is critical for the success of any self-driving car project.'[36]

It is the human component of driving and its unwritten rules that current autonomous technology fails fully to grasp and act on in a manner deemed safe for public use. At present the risks of self-driving cars outweigh the rewards. An article in *The National Law Review* pointed out that statistically there are 9.1 driverless-car accidents per million miles compared to 4.1 for regular vehicles, making autonomous vehicles more than twice as dangerous.[37] It remains to be seen how the promise of autonomous driving evolves over time, but as of now it is an example of where AI is not yet fit for purpose.

NOTES TO CHAPTER 2

1 Costica Dumbrava, 'Artificial Intelligence at EU Borders: Overview of Applications and Key Issues', *European Parliamentary Research Service*, July 2021, p. 2.
2 'Bar Code Scanner', *Encyclopedia.com*, 18 January 2023, https://www.encyclopedia.com/science-and-technology/computers-and-electrical-engineering/electrical-engineering/optical-scanners.
3 Thales Group, 'The Electronic Passport in 2021 and Beyond', 23 February 2023, https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/passport/electronic-passport-trends.
4 UK Border Force, 'About Us', 12 March 2023, https://www.gov.uk/government/organisations/border-force/about.
5 HM Passport Office, 'Biometric Passports and Passport Readers', 22 March 2023, https://www.gov.uk/government/publications/biometric-passports-and-passport-readers/biometric-passports-and-passport-readers.
6 Dominik Malčík and Martin Drahanský, 'Anatomy of Biometric Passports', *Journal of Biomedicine and Biotechnology* 2012:4, Article ID 490362, 18 July 2012, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3433183/pdf/JBB2012-490362.pdf.

7 Yassin Kortli, Jridi Maher, Ayman Al Falou and Mohamed Atri, 'Face Recognition Systems: A Survey', *Sensors* 20:2, 7 January 2020, https://www.mdpi.com/1424-8220/20/2/342/htm.
8 Kortli et al., 'Face Recognition Systems'.
9 Kortli et al., 'Face Recognition Systems'.
10 'About Face ID Advanced Technology: Find out how Face ID helps protect your Information on your iPhone and iPad Pro', *Apple*, 7 April 2023, https://support.apple.com/en-gb/HT208108.
11 Peter Schmallegger, 'The Future of ePassports and Border Crossings', *NXP Semiconductors*, 5 March 2023, https://www.nxp.com/docs/en/white-paper/the-future-of-epassports-and-border-crossings-whitepaper.pdf.
12 Malčík and Drahanský, 'Anatomy of Biometric Passports'.
13 Jim Canham, 'Emerging Technologies can Transform Border Management, but Agencies must Prepare', *WCO News*, 11 January 2023, https://mag.wcoomd.org/magazine/wco-news-82/emerging-technologies-can-transform-border-management-but-agencies-must-prepare/.
14 Autumn Cafiero Giusti, 'E-Gates Ease and Secure International Travel: Kiosks read Passport, use Biometrics to verify Identity', *Secure ID News*, 18 December 2016, https://www.secureidnews.com/news-item/e-gates-ease-and-secure-international-travel/2/.
15 'Automated Fingerprint Identification System Market is likely to reach a Valuation of nearly US$ 85 bn by the End of 2032', *Yahoo! Finance*, 4 October 2022, https://finance.yahoo.com/news/automated-fingerprint-identification-system-market-143000667.html.
16 Australian Government, Department of Immigration and Border Protection, 'Arrivals SmartGate', 19 January 2023, https://web.archive.org/web/20160103192933/http://www.border.gov.au/Trav/Ente/Goin/Arrival/Smartgateor-ePassport.
17 'What was the Real Futurama?', *YouTube*, 13 January 2021, https://www.youtube.com/watch?v=aMSp4d2rOzE.
18 Roland Marchand, 'The Designers go to the Fair II: Norman Bel Geddes, the General Motors "Futurama," and the Visit to the Factory Transformed', *Design Issues* 8:2 (1992), pp. 22–40.
19 P. M. Fotsch, 'The Building of a Superhighway Future at the New York World's Fair', *Cultural Critique* 48 (Spring 2001), pp. 65–97.
20 Bonnie Gringer, 'History of the Autonomous Car', 12 March 2023, https://www.titlemax.com/resources/history-of-the-autonomous-car.

21 Rebecca Fannin, 'Where the Billions spent on Autonomous Vehicles by U.S. and Chinese Giants is heading', *CNBC*, 21 May 2022, https://www.cnbc.com/2022/05/21/why-the-first-autonomous-vehicles-winners-wont-be-in-your-driveway.html.
22 'Elon Musk: Self-driving is way harder than I thought', Lex Fridman Podcast Clips, *YouTube*, 1 January 2022, https://www.youtube.com/watch?v=MyLUiE-XfQI.
23 Frank Landymore, 'Godfather of Self-Driving Cars says the Tech is going Nowhere: "It's an Illusion"', *Futurism*, 10 September 2022, https://futurism.com/self-driving-industry-going-nowhere.
24 'Chassis: The First Suspension System with "Eyes" now sees even Better', *Mercedes-Benz Group Media*, 22 October 2022, https://group-media.mercedes-benz.com/marsMediaSite/en/instance/ko/Chassis-The-first-suspension-system-with-eyes-now-sees-even-better.xhtml?oid=22944501.
25 'Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles', *Society of Automotive Engineers*, 28 March 2023, https://www.sae.org/standards/content/j3016_202104/.
26 Araz Taeihagh and Hazel Si Min Lim, 'Governing Autonomous Vehicles: Emerging Responses for Safety, Liability, Privacy, Cybersecurity, and Industry Risks', *Transport Reviews* 39:1 (2019), pp. 103–28.
27 Hugh Durrant-Whyte and Tim Bailey, 'Simultaneous Localization and Mapping: Part 1', *IEEE Robotics and Automation Magazine* 13:2 (2006), pp. 99–110.
28 Talha Takleh Bin Omar Takleh, Nordin Abu Bakar et al., 'A Brief Survey on SLAM Methods in Autonomous Vehicle', *International Journal of Engineering and Technology (UAE)* 7:4 (2018), pp. 38–43.
29 Lora Kolodny, 'Cruise gets Green Light for Commercial Robotaxi Service in San Francisco', *CNBC*, 2 June 2022, https://www.cnbc.com/2022/06/02/cruise-gets-green-light-for-commercial-robotaxis-in-san-francisco.
30 'Uber Self-Driving Crash: Footage shows Moment before Impact', *BBC News*, 22 March 2018, https://www.bbc.co.uk/news/world-us-canada-43497364.
31 Troy Griggs, 'How a Self-Driving Uber Killed a Pedestrian in Arizona', *The New York Times*, 18 May 2018, https://www.nytimes.com/interactive/2018/03/20/us/self-driving-uber-pedestrian-killed.html.
32 Ryan Randazzo, 'What went wrong with Uber's Volvo in Fatal Crash? Experts shocked by Technology Failure', *AZ Central*, 22 March 2018, https://eu.azcentral.com/story/money/business/tech/2018/03/22/what-went-wrong-uber-volvo-fatal-crash-tempe-technology-failure/446407002/.

33 Mike Isaac and Daisuke Wakabayashi, 'A Lawsuit against Uber highlights the Rush to conquer Driverless Cars', *The New York Times*, 24 February 2017, https://www.nytimes.com/2017/02/24/technology/anthony-levandowski-waymo-uber-google-lawsuit.html.

34 Matthew Sparkes, 'Driverless Cars can be Tricked into Seeing Red Traffic Lights as Green', *New Scientist*, 14 April 2022, https://www.newscientist.com/article/2315634-driverless-cars-can-be-tricked-into-seeing-red-traffic-lights-as-green/.

35 Radhika Madhavan, 'How Self-Driving Cars work: A Simple Overview', *EMERJ*, 3 June 2019, https://emerj.com/ai-sector-overviews/how-self-driving-cars-work/.

36 'Join Us – Software Engineer, Evaluation, Behavior Prediction' [job advertisement], *Waymo*, 19 January 2023, https://web.archive.org/web/20221007194952/https://waymo.com/intl/es/joinus/4358489/.

37 Clifford Law Offices, 'The Dangers of Driverless Cars', *The National Law Review* 11:125 (5 May 2021), https://www.natlawreview.com/article/dangers-driverless-cars.

Chapter 3

Developing a Judaeo-Christian understanding of AI

Chapter 1 considered key milestones in the history of artificial intelligence; Chapter 2 offered two case studies illustrating its benefits and potential pitfalls. Here the focus is on developing an understanding of AI through the lens of Judaeo-Christian teaching. It is important to reiterate the distinction between currently available technology (narrow AI and machine learning) and what is hypothesised (general AI and superintelligence). The following will touch on both, to develop a theistic approach to AI's role and appropriate utility in contemporary life. The first section takes an exegetical approach to relevant texts within Scripture and highlights the uniqueness of humanity within creation; the second seeks to offer a balanced Judaeo-Christian response to AI.

3.1 An ontology of creation

A central theme on the role of human beings within creation is the notion of spiritual transcendence as a means of continual transformation and divinely guided agency. The ontology of the biblical account of creation points to an existence in which there is a creator and a creation that bears the creator's image. Physical matter is not eternal but created *ex nihilo*, through divine intervention. The Judaeo-Christian world view therefore emphasises a creation that not only bears the image of the creator but presents a reality in which the spiritual realm is present in the temporal. Bearing the creator's image enables human beings to enjoy the presence of a divine spirit that elevates life on earth into a renewed metaphysical reality ('The Word became flesh' – John 1.14 NIV). The Judaeo-Christian account of creation therefore rejects the reductionistic belief that human beings are mere contingent biochemical entities, bound to the material, objective world of matter. They are also spiritual beings capable of transcending that world. Hence they are not bonded to operate exclusively by and for computational challenges.

An exegetical approach to some key Old and New Testament texts is helpful in understanding a biblical ontology of creation. Chronologically speaking, the first such passage is Genesis 1.27–28, where 'God created mankind in his own image … God blessed them and said to them, "Be fruitful and increase in number; fill the earth and subdue it"' (NIV). Genesis offers several ontological observations on the centrality of the human being in creation. The first would be that humans are fundamentally distinct from other living creatures. They – male and female – are created in the image of God, or *imago Dei*. From inception they bear the marks and some characteristics of the creator. Although the exact extent or nature of such divine characteristics have been much debated, the essence of what it means to be human is to be an image bearer of God. David Atkinson usefully unpacks the premise, suggesting that: 'what the phrase "image of God" is pointing to is this question: What does it mean to be authentically human?'[1]

One feature of being 'authentically human' is the capacity for self-awareness and profound self-reflection. God represents the ultimate one who is self-aware, and we have the capacity to replicate: 'to be in his image is to be aware of ourselves as his creatures.'[2] Humanity therefore reflects the creator in both the detail of human–God attributes and the context of earthly human existence. No other living creature has this ability. From this dualistic material and metaphysical existence stems a consciousness that compels humans to act, at least in part, according to some inherent perception of a moral law. This does not necessarily imply that human actions always lead to moral outcomes but that, unlike other created beings, humans have an intrinsic ability for moral reasoning that sees beyond short-term action.

There are therefore two key points here. First, since we are created in the image of God, humanity carries dignity – derived from its divine creator. Second, we reflect the character of our creator and hence a perception of the moral law. Consequently, a central part of what it means to be human is this capacity for moral reasoning and for reflecting the dignity and character of God.

An Augustinian approach gives precedence to the *imago Dei* as the relationship between God and the human soul. The power of this relationship not only makes humans unique in their ability to connect and – in Christ – become one with the creator, it also allows them to experience love in the most profound and transformative sense of the word. Agape love is sacrificial love offered by the creator to humanity unconditionally – regardless of any contingent circumstances or human action. In Scripture we learn that God is infinite (Colossians 1.17), immutable (Malachi 3.6), omnipotent (Isaiah 43.13), omniscient (Isaiah 46.9–10), omnipresent (Psalm 139.7–10), holy (Revelation 4.8), righteous (Deuteronomy 32.4), merciful (Romans 9.15–16); and finally, we learn that God is love (1 John 4.8).

In a sermon about God's love, St Augustine of Hippo (AD 354–430) encouraged his congregation to 'Love and do what you will.'[3] This has often been misinterpreted as meaning that it does not matter what one does: as long as one loves God one may do as one pleases. The correct meaning is that God's love relationship towards those who sincerely seek him will transform and guide their lives. The connection and love relationship with the creator is what allows humans ultimately to fulfil their calling – both in this world and beyond. Teodora Prelipcean argues that:

> by complying with ['Love and do what you will'], our thoughts and behaviour will be the beneficial consequences of this love … The supreme virtue and the essence of moral life, love is

> always accompanied by many other human qualities: faith, justice, kindness, altruism, sincerity, understanding, modesty, good will, tolerance, patience, mercy, parsimony, etc.[4]

Love as the supreme human virtue is something gifted to us from God. We are the only vessels capable of receiving love in such form. The transformative power of this intimate agape love is expressed in its fullness in Christ. The Triune God – Father, Son and Holy Spirit – is defined as three persons consubstantially rooted in the perfect relationship of selfless love. In the incarnation, when the Son embraces full humanity, it is this interpersonal selfless love that becomes available to all. It is in him, as the true *imago Dei*, that true humanity is being revealed, and it fundamentally includes the question of love. This is indeed the uniqueness of the Christian faith: in Christ, divine, eternal and selfless interpersonal love is not only revealed as an external reality, it becomes fully available and accessible to every human being. True humanity is therefore both interrelational and fundamentally defined by love.

Those in Christ become partakers in this perfect interpersonal love, being also enabled to manifest – albeit in limited ways – such love. In Augustinian terms, the human being is called to contemplate the divine love and emulate it in his or her own life – to manifest such love towards God, towards the self and towards others. The Christian definition of being human goes beyond being an isolated, individualistic rational self: René Descartes' mind–body dualism within the concept of a 'Cartesian subject' is just one limited definition of being human.[5] In the Christian tradition, the human subject is defined within the logic of the perfect interpersonal love of the triune God. The patristic *Amo ergo sum*, 'I love therefore I am', is in this respect superior to the mere 'I think, therefore I am', *Cogito ergo sum*.

Paul repeatedly emphasises this love process of transformation and renewal into the image and likeness of Christ (Ephesians 4.24; 2 Corinthians 5.17; Colossians 3.10). For Paul, Christ is the redemptive saviour who reconciles a humanity separated by sin from the creator. This is achieved by grace through faith in Jesus Christ (Ephesians 2.8–9).

The Pauline epistles mark a call to humanity as bearers of the image of God to be transformed in Christ and enter into a new existence in which individuals are called to 'put off your old self, which is being corrupted by its deceitful desires; to be made new in the attitude of your minds' (Ephesians 4.22–23 NIV). This is made possible through Christ, who not only represents the perfect image of God but is God incarnate (John 1.14). Christ is 'the image of the invisible God, the firstborn over all creation' (Colossians 1.15 NIV). He is the second Adam, the true human (1 Corinthians 15.45–49). His eternal presence precedes and is not dependent on his physical embodiment. He is both the means of reconciliation and the end itself; that is, eternal life. Through Christ, all things were created and are sustained (Colossians 1.17). The creator's aim, therefore, is reconciliation with a fallen humanity, ultimately achieved at the cross and resurrection. The eternal has entered the temporal, making what is temporal eternal. To be truly human means to be able to imagine, and access, what is eternal.

Reflecting on Colossians 1.17, Professor Robert Wall writes: 'Though the material effects of sin and fallenness remain all too evident, Paul can claim that the Creator's goal has already been realised through Christ and is already being demonstrated in the life of a new creation, the church.'[6] Paul then introduces the concept of Christ as the head of the body, which is the Church (Colossians 1.18). The emphasis is again on the importance of humanity as image bearers of God and the relational dimension that emerges as a result. N. T. Wright argues that: 'the metaphor of a human body is utterly appropriate to express not only mutual interdependence but also, as here, an organic and dependent relation to Christ himself.'[7] The image of a body evokes a certain degree of unity and, indeed, interconnectedness – Paul details the more practical consequences of being part of the body of Christ in 1 Corinthians 12.12–27.

What are the implications for humans as image-bearers of God? A holistic Judaeo-Christian picture depicts a humanity capable of a transformative – and redemptive – relationship with God, with the self and the other, and eventually with the whole of creation. Humanity is designed for communion with the creator, which distinguishes it from other living beings. It has the capacity not only to go beyond its contingent reality but to imagine eternity and find ways towards it. The relational dimension of Christianity expressed in the form of an ontological contingency was also emphasised by Pope John Paul II: 'Man is an autonomous subject. He is the source of his own actions, while maintaining the characteristics of dependence on God, the Creator.'[8] A person transcends the physical limitations of their body in their freely choosing God.[9] They are *capax Dei* – capable of knowing God and receiving spiritual gifts from God. Authority is therefore also derived from above. Wright points out that it is precisely the *imago Dei* that gives humanity stewardship over creation: 'When humans praise God, they ought to realise that they are doing so as representatives of the whole world.'[10] They are the true 'royal priesthood' (1 Peter 2.9 NIV).

3.2 Judaeo-Christian teaching and AI: towards a human-centric AI?

As argued above, Scripture places the human being as both body and soul created for transformative relationship with the divine and with others. The relationship with and likeness to the creator give humans a unique position of dignity, authority and stewardship over the rest of creation. Therefore what might a Judaeo-Christian approach to the development of AI look like?

It is useful to start by noting that even among secular thinkers there is much debate on the nature of the mind and artificial intelligence. David Bolter has written extensively on the topic and views computer intelligence only as a metaphor for human intelligence:

> The artificial intelligence specialists have, I think, gone too far. The computer is a mirror of human nature … it is not a perfect mirror; it affects and perhaps distorts our gaze, magnifying certain human capacities (those most easily characterized as 'information processing') and diminishing others.[11]

Others, such as Herbert Roitblat, recognise that artificial intelligence and human intelligence have very little in common:

> The emphasis of intelligence testing and computational approaches to intelligence has been on well-structured and formal problems … But we humans are creative, irrational, and inconsistent … We do sometimes behave like computers, but more often, we are sloppy and inconsistent.[12]

There are also points of disagreement among faith communities when it comes to technology and the use of AI. Much of this stems from the orthodox biblical view of humans as the sole transcendent beings within creation. This implies they are not capable of creating through their own means entities that can equal or surpass them in the spiritual dimension; that is, in the God-given attributes that make them truly human. Only a transcendent God can do that. Regardless of how autonomous or sentient an AI may appear, the depths of its essence will always remain in the temporal, computational dimension. AI and humanity cannot and will never share the same essence and substance.

A transhumanist view of the future challenges this by claiming that humans can be profoundly spiritually transformed and improved through AI. The more contentious issue here is the belief that a superintelligent AI will invariably possess a soul and conscience. This, transhumanists argue, has deep implications for the gospel message and salvation: humanity now has the responsibility to bring salvation to all people – and robots. Christopher Benek, a Florida pastor who describes himself as a 'techno-theologian, futurist, ethicist, Christian Transhumanist',[13] has said: 'I don't see Christ's redemption limited to human beings … It's redemption to all of creation, even AI. If AI is autonomous, then we should encourage it to participate in Christ's redemptive purposes in the world.'[14] This is a confusion of category. Dr Beth Singler of the Faraday Institute for Science and Religion asks:

> If we create AI beings in our image, will they have a relationship with us and a relationship with God? Will they, in fact, be spiritual beings? … How will they [people of faith] respond when a robot turns up at a church, mosque, or synagogue and says, 'I believe'?[15]

One church in Poland already uses a robot priest named SanTO to help impart biblical knowledge to its parishioners. Some locals say it's 'a bit like Catholic Alexa', while others put it more bluntly: 'It has no soul, it is not a person.'[16]

The transhumanist world view raises a number of theological issues, or fallacies, some more obvious than others. First, the transhumanist search for an ultimate cyborg-type superintelligence that is part-human, part-machine has been labelled by Professor John Lennox a 'flawed narrative'. If there is a God, that God is the superintelligence that has always existed: 'He is not an End Product. He is the Producer.'[17] God's omnipotent ability to create in his own image does not imply that, by extension, humans have the ability to create in theirs: 'You have made them a little lower than the angels and crowned them with glory and honour. You made them rulers over the works of your hands' (Psalm 8.5–6 NIV). Key here is 'rulers': human beings are rulers – or caretakers – of creation, not its creators. To entertain the idea of a sentient AI of equal or comparable spiritual and emotional value to a human is simply illogical from a theological perspective. It is impossible to assign any innate spiritual value to a man-made machine.

The second issue lies at the intersection of false idolatry and anthropomorphisation – the Eliza effect mentioned in Chapter 1. There are substantial differences between appearing conscious – via deep learning, for instance – and being conscious. God may choose to use technology as a means to reach humans and advance his will on creation; this does not imply that the technology itself carries any innate form of spiritual capability or value. While God's provision of technology can be spiritually strengthening, any assumption of spiritual capacity in the technology itself may amount to anthropomorphisation or even idolatry. A superintelligent AI demonstrates no more consciousness by walking into church and saying 'I believe' than a cash machine does by displaying 'Thank you for your transaction. Have a nice day.'

The third and rather obvious issue is that artificial intelligence lacks the highly sophisticated biochemical build of the human body. Therefore it does not possess the biological–intellectual–spiritual interconnectedness that plays such a central role in human development, thought processing and decision-making. Human behaviour and human intelligence are cumulatively products of this complex biochemical–spiritual makeup. The tech author James Hoskins sums up the problem:

> Consciousness includes much more than just a quantitative measure of intelligence. It includes qualitative experiences such as subjective awareness, understanding, intentionality, and the unity of one's self-identity. Even if a computer were intelligent enough to make it effectively appear as though it was having qualitative experiences – such as if it acted like it was in pain, or in love, for example – we still could not be sure it was truly conscious.[18]

This reaffirms that through the essence of their being, humans possess a unique gift that allows them to surpass their biochemical build-up and enter a relationship with the divine – a relationship that entails continual growth and the transformation of the created being into the image and likeness of the creator. Regardless how complex a superintelligent machine might appear, its essence will always remain of the same matter as any other electronic device. There is always a man-given – thus contingent – material location a soulless machine will never be able to transcend.

This results in a decoupling of intelligence and consciousness – the two are not the same. Lennox points out that in Genesis an omnipotent God, through his spirit, 'linked intelligence and consciousness in one being'.[19] Therefore consciousness does not necessarily rely on physical matter, rather it is an attribute divinely gifted, which makes it impossible for humans to impart consciousness to other created machines. This comes back to the reality that humans created in the image of the triune God find themselves in a relationship that, to a greater or lesser extent, informs their thoughts and guides their actions. Humanity is not alone.

Love, of course, is the epitome. To be in a relationship means to know and experience love. As fundamentally relational creatures, humans were created for love. Professor Stephen Williams says that: 'humans are formed by God in the very core of their being to be recipients and givers of love.'[20] It is through love that humans are given the greatest commandment to 'Love the Lord God with all your heart and with all your soul and with all your mind and with all your strength' and 'Love your neighbour as yourself' (Mark 12.30–31 NIV). God's image and providence prompt a relationship that spills over into love of self and love of neighbour – both created as image bearers of the divine.

Does this imply that technology or AI represent evil? Not at all. As man-made things they are only as 'good' or 'evil' as the programmers create or allow them to be. Certain Christian denominations have already issued guiding statements on the adoption and use of AI. Most notably, in the USA, the Ethics and Religious Liberty Commission of the Southern Baptist Convention has published a statement of principles. It embodies much of what has been said thus far and could serve as a robust template for what might constitute an evangelical approach to AI. The following are some of the most relevant extracts:

> We deny that any part of creation, including any form of technology, should ever be used to usurp or subvert the dominion and stewardship which has been entrusted solely to humanity by God; nor should technology be assigned a level of human identity, worth, dignity, or moral agency.
>
> We affirm that the development of AI is a demonstration of the unique creative abilities of human beings. When AI is employed in accordance with God's moral will, it is an example of man's obedience to the divine command to steward creation and to honor Him. We believe in innovation for the glory of God, the sake of human flourishing, and the love of neighbor. While we acknowledge the reality of the Fall and its consequences on human nature and human innovation, technology can be used in society to uphold human dignity. As a part of our God-given creative nature, human beings should develop and harness technology in ways that lead to greater flourishing and the alleviation of human suffering.
>
> We deny that AI will make us more or less human, or that AI will ever obtain a coequal level of worth, dignity, or value to image-bearers. Future advancements in AI will not ultimately fulfill our longings for a perfect world. While we are not able to comprehend or know the future, we do not fear what is to come because we know that God is omniscient and that nothing we create will be able to thwart His redemptive plan for creation or to supplant humanity as His image-bearers.[21]

The Catholic Church has also expressed views on the subject in calling for AI to be used in the service of humanity.[22] In 2020, the Pontifical Academy for Life co-signed, with Microsoft, IBM, the FAO and the Italian Ministry of Innovation, the 'Rome Call for AI Ethics'. This posited six core principles to guide the future development of AI:

> Transparency: AI systems must be understandable to all.
>
> Inclusion: these systems must not discriminate against anyone because every human being has equal dignity.
>
> Accountability: there must always be someone who takes responsibility for what a machine does.
>
> Impartiality: AI systems must not follow or create biases.
>
> Reliability: AI must be reliable.
>
> Security and Privacy: these systems must be secure and respect the privacy of users.[23]

In November 2021, at the 41st Session of the UNESCO General Conference, Cardinal Pietro Parolin said:

> For the Holy See, the principle that not everything that is technically possible or viable is thereby ethically acceptable remains ever valid. In order to be able to speak correctly of an ethics of artificial intelligence, it will therefore be necessary that the development of every algorithm always draws on an ethical vision, 'algor-ethics'.[24]

The Church of England has yet to publish an official statement on the adoption and use of AI. However, an advisory paper, 'Big Tech', recommends that technology companies strive to create new technologies with:

- a commitment to verifiable transparency;
- a commitment to promote human-centred design;
- a commitment to enable the flourishing of children and other vulnerable groups;
- commitment to foster a tech ecosystem that serves the common good.[25]

Broadly speaking, all Christian denominations emphasise a human-centric adoption of AI. This offers a useful segue into developing a Christian approach to AI, which will focus on narrow AI. Again, for present purposes narrow AI is defined as technology that is currently available, incorporating traditional computational devices as well as more recent machine-learning capabilities and the use of artificial neural networks.

Judaeo-Christian teaching views human ingenuity and creativity as a gift from God that forms a part of what it means to be made in the image of God. Technological development can therefore be seen as part of the divine imperative of stewardship, value and wealth creation. Technology is, in effect, an extension of mechanisms and objects created by humans and is thus profoundly intertwined with core human traits and abilities, such as innovation, creativity and purposefulness – all cumulatively reflecting God's divine character.

Advancements in technology can therefore be viewed as part of the wider mandate and pursuit of work itself. In Genesis 1.28, where humanity is called to rule and subdue the earth, the stewardship commandment implies the use of all divine gifts bestowed on humanity, including intelligence and creativity. This was integrated in the Ethics and Religious Liberty Commission of the Southern Baptist Convention statement above, where AI 'is an example of man's obedience to the divine command to steward creation and to honor Him'.[26] There are also relevant passages from Scripture: Proverbs 18.15 states that 'Intelligent people are always ready to learn. Their ears are open for knowledge' (NLT); King David in 2 Chronicles 2.12 was 'endowed with intelligence and discernment' (NIV). Indeed, it is only reasonable to assume that Jesus himself must have used practical intelligence and creativity throughout his work as a carpenter (Mark 6.3). It can be argued that Jesus was creating technological objects. Intelligence and creativity therefore permeate all of creation, and Judaeo-Christian teaching ultimately reveals them as part of God's creative character, and his purpose to use them for good. Richard Turnbull has written that: 'God's action in creation is the supreme creative act, reflected in both human nature and human purpose.'[27]

Where would problems arise when it comes to narrow AI? To start with, Scripture is clear that all forms of idolatry are to be shunned: 'You shall not make for yourself an image in the form of anything in heaven above or on the earth beneath or in the waters below. You shall not bow down to them or worship them' (Exodus 20.4–5 NIV), which would represent an extreme form of adoration for something or someone that directly competes or supersedes the adoration of God.

This can lead to where narrow AI ceases to contribute to human flourishing and becomes detrimental or destructive. According to Pope John Paul II, we are called to:

> use science and technology in a full and constructive way, while recognizing that the findings of science always have to be evaluated in the light of the centrality of the human person, of the common good and of the inner purpose of creation.[28]

Pope Francis has called the internet a 'gift from God', yet urges caution and warns against an overreliance on technology – too much information can cause 'mental pollution' and harm interpersonal relationships. He has also said that online media 'can stop people from learning how to live wisely, to think deeply and to love generously'.[29]

Examples abound of this paradoxical dualism presented by current use of narrow AI – including the case studies above on the benefits of biometric passports and risks of autonomous vehicles. There are also deep repercussions in such areas as social media platforms: at what point do they stop benefitting users and start becoming detrimental? On the one hand, the sharing of information freely is a huge development (pictures, videos, communication with friends and family); on the other, the veil of anonymity has led to cyberbullying, blackmail, threats and verbal aggression. Unfortunately, young people seem worst affected, which in the USA has seen suicide rates among teenagers rise by 57.4 per cent between 2007 and 2018 – an increase that seems largely driven by reliance on and daily use of social media.[30]

Other associated risks of narrow AI may be more subtle and harder to identify. In the work environment, online videoconferencing and Covid-driven remote working can provide speed and accessibility, but sometimes at the expense of human interaction. Computer games provide endless hours of entertainment for youngsters while isolating them from their peers. Online shopping algorithms often use machine learning to present and reinforce unhealthy habits, such as gambling or overspending. There is of course the issue of privacy, which despite its prominence is not yet widely understood. More work needs to be done until all online users learn that the data they generate becomes the currency with which they pay for 'free' services. Privacy becomes less of a moral concern once users know how online data is collected and monetised.

The problems surrounding humanity's embrace of technology are therefore multiple and vary in severity, while some solutions are clearer than others. A Judaeo-Christian approach to narrow AI requires discernment, moderation and an emphasis on the centrality of the human being at the core of creation under God. The use and spread of narrow AI must also be viewed within the context of the fall (Genesis 3), where mankind's sin pervades all that is produced. AI is no exception to this precarious environment. Used as part of God's intended purpose in creation, AI contributes to human flourishing; but in the hands of fallen human beings, the possibility of misuse is very real. AI is fundamentally a product of human creativity and as such ought to be embraced in a prudent manner that directs its contributions towards human thriving and stewardship of the rest of creation.

Notes to Chapter 3

1 David Atkinson, *The Message of Genesis 1—11: The Dawn of Creation* (The Bible Speaks Today), Westmont, IL: InterVarsity Press, 1990, p. 36.

2 Atkinson, *The Message of Genesis 1—11*, p. 37.

3 Christian History Institute, 'Augustine's Love Sermon', 12 May 2023, https://christianhistoryinstitute.org/study/module/augustine.

4 Teodora Prelipcean, 'Saint Augustine – The Apologist of Love', *Procedia – Social and Behavioral Sciences* 149 (2014), pp. 765–71.

5 'Dualism', *Stanford Encyclopedia of Philosophy*, 5 April 2023, https://plato.stanford.edu/entries/dualism/.

6 Robert W. Wall, J. Paul Sampley, Richard B. Hays and N. T. Wright, *The New Interpreter's Bible Commentary, Vol. IX: Acts; Introduction to Epistolary Literature; Romans; 1 & 2 Corinthians; Galatians*, Nashville, TN: Abingdon Press, 2002, p. 62.

7 N. T. Wright, *Colossians and Philemon* – Tyndale New Testament Commentaries, Westmont: Inter-Varsity Press, 1987, p. 74.

8 Pope John Paul II, 'Humans are Created in the Image of God', *INTERS: Interdisciplinary Encyclopedia of Religion and Science*, 9 April 1986, https://inters.org/John-Paul-II-Catechesis-Image-God.

9 Mark K. Spencer, 'Perceiving the Image of God in the Whole Human Person', *The Saint Anselm Journal* 13:2 (Spring 2018), pp. 13–14.

10 'N. T. Wright on what it means to be an Image Bearer', *YouTube*, 17 April 2023, https://www.youtube.com/watch?v=yp-Ku-_ekAY.

11 J. David Bolter, 'Artificial Intelligence', *Daedalus* 113:3 (Spring 1984), pp. 1–18.

12 Herbert L. Roitblat, 'AI is no Match for the Quirks of Human Intelligence', *The MIT Press Reader*, 4 November 2022, https://thereader.mitpress.mit.edu/ai-insight-problems-quirks-human-intelligence/; Roitblat is also the author of *Algorithms Are Not Enough: Creating General Artificial Intelligence*, Cambridge, MA: MIT Press, 2020.

13 Christopher Benek, 'About', 20 March 2023, https://www.christopherbenek.com/about/.

14 Micah Redding, 'Why I became a Christian Transhumanist', *Vice*, 18 December, 2022, https://www.vice.com/en/article/9akxm3/why-i-became-a-christian-transhumanist.

15 'Robots and Religion', *Christian Evidence*, 15 April 2023, https://christianevidence.org/2017/01/16/robots_and_religion/.

16 Greig R. Hill, '"Meet Your Maker": The Robot Priests taking the World by Storm', *Medium*, 17 December 2022, https://medium.com/in-our-times/meet-your-maker-the-robot-priests-taking-the-world-by-storm-32b2e398383.

17 John Lennox, *2084: Artificial Intelligence and the Future of Humanity*, Grand Rapids, MI: Zondervan, 2020, p. 157.
18 James Hoskins, 'Digital Souls: What should Christians believe about Artificial Intelligence?', *Christian Research Journal* 39:2 (2016), p. 5.
19 Lennox, *2084*, p. 127.
20 Stephen N. Williams, 'What is it to be a Person?', in John Wyatt and Stephen N. Williams (eds), *The Robot Will See You Now: Artificial Intelligence and the Christian Faith*, London: SPCK, 2021, p. 102.
21 The Ethics and Religious Liberty Commission (ERLC), 'Artificial Intelligence: An Evangelical Statement of Principles', 12 April 2023, https://erlc.com/resource-library/statements/artificial-intelligence-an-evangelical-statement-of-principles/.
22 Jonah McKeown, 'Sentient AI? Here's what the Catholic Church says about Artificial Intelligence', *Catholic News Agency*, 18 March 2023, https://www.catholicnewsagency.com/news/251552/sentient-ai-heres-what-the-catholic-church-says-about-artificial-intelligence.
23 'Rome Call for AI Ethics', https://www.romecall.org/wp-content/uploads/2022/03/RomeCall_Paper_web.pdf.
24 Aaron Humphriss, 'Artificial Intelligence and the Catholic Church', *The Catholic Union of Great Britain*, 25 March 2023, https://catholicunion.org.uk/2022/01/ai/.
25 'Big Tech: The Policy of the Church of England National Investing Bodies and the Advice of the Church of England Ethical Investment Advisory Group', *The Church of England*, 17 May 2023, https://www.churchofengland.org/sites/default/files/2022-09/big-tech-policy-and-advice.pdf.
26 Ethics and Religious Liberty Commission, 'Artificial Intelligence'.
27 Richard Turnbull, *Work as Enterprise*, Oxford: The Centre for Enterprise, Markets and Ethics, 2020, p. 35.
28 'Common Declaration on Environmental Ethics – Common Declaration of John Paul II and the Ecumenical Patriarch His Holiness Bartholomew', *The Vatican*, 5 May 2023, https://www.vatican.va/content/john-paul-ii/en/speeches/2002/june/documents/hf_jp-ii_spe_20020610_venice-declaration.html.
29 Issie Lapowsky, 'What does the Pope think about Technology? #It'sComplicated', *Wired*, 11 May 2023, https://www.wired.com/2015/09/pope-francis-technology/.
30 Sally C. Curtin, 'State Suicide Rates among Adolescents and Young Adults aged 10–24: United States, 2000–2018', *National Vital Statistics Reports* 69:11 (12 April 2023), https://www.cdc.gov/nchs/data/nvsr/nvsr69/NVSR-69-11-508.pdf.

Conclusion

The trajectory of its future development cannot be known but it is clear that artificial intelligence will become increasingly embedded in the home and beyond. The complexities of AI mean it will have varied impact across different sectors of the economy and on their respective pathways of growth or decline. Jobs that are repetitive in nature and/or work towards binary outcomes are particularly at risk. Managing transformational AI will therefore require a degree of wisdom, understanding and individual responsibility on the part of programmers and end-users. At present there is a combined effort in establishing a permanent feedback loop that identifies and rectifies mistakes as the program is used (ChatGPT, for instance, utilises such a system).

The prospect of superintelligent AI raises ontological questions about what it means to be human. Judaeo-Christian teaching emphasises the implications of human beings as bearers of the *imago Dei*, and the role and responsibility this carries vis-à-vis the rest of creation. It therefore offers a unique perspective in developing AI in which the human being remains central.

Chapter 1 touched on key milestones in AI history, highlighting the gradual advancement from basic computational AI to such new technologies as machine learning, pattern recognition and deep learning. Chapter 2 presented two case studies: an analysis of AI-led biometric identification, which brought significant efficiency and security benefits to points of transit; and the development of autonomous vehicles which, though innovative, are as yet unsafe for widespread public use. Chapter 3 considered a Judaeo-Christian response to the adoption and use of AI. It emphasised the reality of humanity as bearing the image of the divine and the implications this has for sentience as a uniquely human attribute.

This concluding chapter offers three guiding principles for AI's adoption and use.

First and foremost, AI has to be utilised and designed in a manner that holistically benefits humanity. Stuart Russell and Peter Norvig have written: 'Given that AI is a powerful technology, we have a moral obligation to use it well, to promote the positive aspects and avoid or mitigate the negative ones.'[1] This is where early adopters and developers of new integrated AI technologies need to dig deeper. The goal should be societal value in forms that enrich relationships and communities, and support economic growth leading towards alleviation of poverty. From the Judaeo-Christian perspective, recognising what distinguishes humans from the rest of creation can act as a guide and catalyst for discerning a right path at a time when the detailed ramifications of AI remain unclear.

The second principle is to ensure that new technologies that can have an impact on the general population are rooted in a set of underlying ethical values. These would generally include such widely held beliefs as human dignity, the respect and protection of life, freedom of expression, fairness in the treatment of others, non-maleficence and justice. The integration of moral values into the AI code should come through, and in some cases from, the computer developers themselves. Developers have a moral duty to ensure to the best of their ability that harmful, unintended consequences are curtailed, and safety systems in place when things go wrong. Some tech companies have already taken concrete steps to implement ethics into their own AI products, and have helped others follow suit. IBM, for instance, launched AI Fairness 360, which is an open-source toolkit available to all developers.[2]

The third principle is that the regulatory framework surrounding AI should focus on the tenets of AI design rather than the intricacies. Though built on a set of principles, the process of innovation must be protected, and corrected at the point at which it fails – innovation should be welcomed with cautious optimism, not cynical suspicion. Adam Thierer, a technology policy expert at R Street Institute, argues against the excessive use of the 'precautionary principle' and points out that: 'If public policy is guided at every turn by fear of hypothetical worst-case scenarios and the precautionary mindset, then innovation becomes less likely.'[3] Policymakers must strive to establish an environment that protects and harnesses the innovative spirit of technology entrepreneurs.

UK legislative proposals and initiatives affecting AI are currently embryonic. A report by the House of Lords Select Committee on Artificial Intelligence acknowledged in 2018 that while 'AI-specific regulation is not appropriate at this stage', the government should create an AI Code that serves as a shared ethical framework.[4] Such a framework would be based on five key principles:

- Artificial intelligence should be developed for the common good and benefit of humanity.
- Artificial intelligence should operate on principles of intelligibility and fairness.
- Artificial intelligence should not be used to diminish the data rights or privacy of individuals, families or communities.
- All citizens have the right to be educated to enable them to flourish mentally, emotionally and economically alongside artificial intelligence.
- The autonomous power to hurt, destroy or deceive human beings should never be vested in artificial intelligence.[5]

In 2020, a subsequent paper by the House of Lords Liaison Committee argued that the Centre for Data Ethics and Innovation within the Department for Science, Innovation and Technology:

> should establish and publish national standards for the ethical development and deployment of AI. National standards will provide an ingrained approach to ethical AI, and ensure consistency and clarity on the practical standards expected for the companies developing AI, the businesses applying AI, and the consumers using AI. These standards should consist of two frameworks, one for the ethical development of AI, including issues of prejudice and bias, and the other for the ethical use of AI by policymakers and businesses. These two frameworks should reflect the different risks and considerations at each stage of AI use.[6]

Vague and broad statements in the development of policy raise additional challenges, not least in defining boundaries of what is deemed ethical by programmers and policymakers. The challenge here is to ensure that regulation maintains a dualistic purpose: pro-innovation and pro-growth, but upholding a basic standard of moral guidance and user security. The innovation-versus-safety dichotomy is not necessarily mutually exclusive. One way to approach this is for regulation to be transparent and coordinated at nascent stages of the development process. Equipping programmers with a clear set of guidelines from the onset substantially reduces the chances of undesirable outcomes. The innovative process can therefore be allowed to develop within the framework rather than be constrained by it (attempts at the latter often occur at an advanced stage of development, when any damage has already been done). The regulatory framework thus needs to ask fundamental questions, such as: What is the purpose of this system? How does it operate and what could be the unintended consequences? How does it integrate a threshold of ethical and safety standards?

Regardless of the exacting regulatory framework, artificial intelligence will be an increasingly potent force in automating and transforming private and public life. AI – and generative AI in particular – will continue to produce the unexpected in solving current problems as well as create new ones. The Judaeo-Christian perspective highlights the centrality of humanity within the rest of creation, and the cosmic order found in Genesis may offer a guiding pathway into the role and scope of AI. A robust approach to AI that challenges its development from the foundational stages, and takes into consideration its aim, nature and purpose in relation to individuals and wider society, leaves us in a more optimistic position to expect positive outcomes in the years ahead.

Notes to Conclusion

1 Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, Global Edition, Harlow: Pearson, 2021, p. 1037.
2 'AI Fairness 360', *IBM Research, Trusted AI*, 20 April 2023, https://www.ibm.com/opensource/open/projects/ai-fairness-360/.
3 Adam Thierer, *Permissionless Innovation: The Continuing Case for Comprehensive Technological Freedom*, Arlington, VA: Mercatus Center at George Mason University, 2014, p. 2.
4 House of Lords, Select Committee on Artificial Intelligence, *AI in the UK: Ready, Willing and Able?*, 16 April 2018, p. 7.
5 House of Lords, *AI in the UK*, p. 125.
6 House of Lords, Liaison Committee, *AI in the UK: No Room for Complacency*, 18 December 2020, p. 11.

APPENDIX

Generative AI: Poem about consciousness (by ChatGPT)

Deep within us all, there lies a spark
A light that flickers, in the dark
A force that drives us, day and night
A consciousness that guides us, towards the light

It's the voice that whispers, in our ears
The one that calms our deepest fears
It's the feeling that tugs, at our heartstrings
The one that tells us, what truly brings

Consciousness, the essence of our being
A force that's infinite, all-seeing
It's the power that drives our every thought
The one that connects us, to what we sought

It's the awareness, of our very soul
The one that makes us, feel whole
It's the spark of life, that's within us all
The one that answers, every call

Consciousness, the source of all creation
The one that fuels our imagination
It's the essence of our humanity
The one that brings us, to our divinity

It's the force that binds, us all together
The one that makes us, feel forever
Consciousness, the gift that we possess
The one that makes us, truly blessed.

Generative AI – Poem about war (by ChatGPT)

A dark cloud looms, a storm of steel and fire,
A clamor of drums, a chorus of desire,
The pounding of boots, the clashing of swords,
The cries of the wounded, the dying, the horde.

Men and women, torn from their homes,
Thrown into battle, to fight for unknown goals,
The colors they bear, the flag they wave,
Blinded by loyalty, to the cause they crave.

Bombs explode, buildings crumble,
Smoke and dust, a world in rubble,
Families torn apart, children left alone,
A land in ruins, a heart of stone.

The enemy, once a stranger, now a foe,
A demon to vanquish, a debt to owe,
The hate that burns, the vengeance sought,
A never-ending cycle, a battle fought.

Yet in the midst of chaos, there is hope,
A glimmer of light, a way to cope,
A hand extended, a heart to heal,
A chance for peace, a new ideal.

For war is not inevitable, it is a choice,
A path we can avoid, a different voice,
Let us lay down our arms, let us seek harmony,
Let us build a world, where war is but a memory.

List of CEME publications

Richard Turnbull, *Quaker Capitalism: Lessons for Today*, 2014.

Edward Carter, *God and Enterprise*, 2016.

Richard Turnbull (ed.), *The Challenge of Social Welfare: Seeking a New Consensus*, 2016.

Richard Turnbull, *The Moral Case for Asset Management* (jointly with New City Initiative), 2016.

Martin Schlag, *Business in Catholic Social Thought*, 2016.

Andrei Rogobete, *Ethics in Global Business*, 2016.

Ben Cooper, *The Economics of the Hebrew Scriptures*, 2017.

Lyndon Drake, *Capital Markets for the Good of Society*, 2017.

Richard Turnbull and Tim Weinhold (eds), *Making Capitalism Work for Everyone*, Vol. 1, 2017.

Richard Turnbull and Tim Weinhold (eds), *Making Capitalism Work for Everyone*, Vol. 2, 2017.

Richard Turnbull, *Understanding the Common Good*, 2017.

Andrei Rogobete, *The Challenges of Migration*, 2018.

Steven Morris, *Enterprise and Entrepreneurship: Doing Good Through the Local Church*, 2018.

Richard Turnbull, *Work as Enterprise: Recovering a Theology of Work*, 2019.

Edward Carter, *God and Competition: Towards a Positive Theology of Competitive Behaviour*, 2019.

Steven Morris, *The Business of God*, 2019.

Andrew Hartropp, *Corporate Executive Remuneration*, 2019.

Richard Turnbull (ed.), *The Economic and Social Teaching of the Hebrew Scriptures*, 2020.

Andrei Rogobete, *The UK Savings Crisis*, 2020.

Barbara Ridpath, *Ethics and Economics: Economics as Servant or Master?*, 2021.

Philip Booth, Kaetana Numa, Stephen Nakrosis, Richard Turnbull, *Government Debt: A Neglected Theme of Catholic Social Teaching*, 2021.

Richard Turnbull (ed.), *The Value of Business: Results of Polling*, 2022.

John Kroencke, *Private Planning and the Great Estates: Lessons from London*, 2022.

Andrei Rogobete, *The Challenge of Artificial Intelligence: Responsibly Unlocking the Potential of AI*, 2023.