The Church of Scotland

Faith Impact Forum

Society, Religion and Technology

Artificial Intelligence (AI)
Opportunity and Challenge for the Church

May 2021

Contents

Artificial Intelligence (AI)

Opportunity and Challenge for the Church

1. Introduction

1.1 Internet-based new technologies are rapidly changing the way we live and have become a remarkable lifeline to people socially isolated during the covid-19 pandemic. Machines with artificial intelligence (AI), at the heart of search engines, voice assistants and speech and face recognition on mobile phones, are powering much of this revolution. AI is used in business and government administration, including the assessment of social security, employment and loan applications. AI algorithms also underpin some impressive innovations in healthcare, assisting doctors to diagnose illness, discover new drugs, read medical images and use robots in surgical operations.

1.2 The Church has many reasons to celebrate and embrace these technologies, but some difficult questions arise about the implementation of the digital revolution. Is our privacy adequately protected by the large tech companies who collect and profit from the personal data of everyone who uses the internet? How much awareness is there amongst users of social media platforms of how those platforms function and monetise them, as users, selling and manipulating their attention for the sake of advertising revenue? How much human supervision is required when life-changing decisions about our health or employment are based on outputs from computers running AI algorithms? Who is held responsible when unfair or damaging decisions are made and do we really understand what these algorithms do? How safe are driverless cars, and who is responsible if things go wrong? How do we respond to the deployment of autonomous weapons or the use of sophisticated AI surveillance tools in policing and the persecution of religious groups by governments? Should we be fearful that runaway technology will move beyond human control? And, ultimately, how does increasing reliance on technology and AI implementations affect who we are as relational human beings, made in the image of God?

1.3 A broad definition of AI is when machines perform tasks which, in the past, only people could do. Algorithms are the sets of mathematical operations that enable a computer to perform these tasks. These can include the rapid analysis of huge data sets and the operation of robots, autonomous cars and drones. In this report we focus on machine learning algorithms, such as those embedded in voice assistants, search engines and speech and face recognition systems. Machine learning is an area of AI which harnesses the ability of computers to rapidly perform repetitive calculations and to recognise patterns and rules found in large sets of data. Artificial neural networks, a development inspired by the working of the human brain, perform "deep learning" to process extremely large amounts of data. These algorithms give machines the autonomy to recognise novel patterns in the data without having to follow pre-programmed human instructions, and (crucially) to act on that new learning.

1.4 The smartness of neural network algorithms was dramatically confirmed in 2016 when an AI programme known as Deep Mind learned the rules of the ancient Asian board game GO by analysing the moves made by champion players in historic matches and went on to outplay the world's best professional GO player. This remarkable milestone in machine learning was soon followed by the AlphaFold2 project, where the colossal processing power of Deep Mind was applied to a longstanding problem in biology: how to predict the three-dimensional folded structure of a protein from a given sequence of amino acids. A breakthrough was achieved in 2020 when Deep Mind took only a few hours to predict the structure of a given protein, completing a task that had previously required several years of laboratory work. This application of AI opens up many exciting new directions of research into the causes and treatments of disease.

1.5 The recent successes of machine learning, tackling problems in chemistry, biology and especially astronomy have been made possible by the increasing power of computer hardware, the creativity of computer scientists developing powerful algorithms and crucially, the availability of very large sets of digital data.

1.6 AI can be applied to almost any task involving big data. When AI methods are used to analyse the data being compiled on all aspects of the lives of billions of people across the world – those who take part in social media and the digital economy – it brings great benefits but can also cause immense harm and be a challenge to human rights.

1.7 In this report we focus on two areas of concern. How do we know that our personal data is used ethically and kept safe and private when we share information on Facebook, Google, Amazon, and the like? And how can the Church best engage with tools using AI for worship and mission when robots and chatbots can plausibly mimic humans, challenging us to reflect on what it means to be a person made in the image of God?

2. Big Data

2.1 The lifeblood of AI is big data, but when these are the personal data of the billions of people who use the internet every day, issues arise about how the information is used. Face recognition is a useful tool to unlock smartphones- but when employed for population surveillance by police forces around the world, the picture changes.

2.2 We share our personal data with every web page we visit, every time we shop online, each internet search, each "like" or "share" on social media, every question offered to a digital personal assistant or as our phones track our location. Digitised data from entire libraries, countless emails and social media content are available to AI developers. One of the largest collections of data, GPT-3 (Generative Pre-trained Transformer version 3 developed by the US company Open-AI) has access to almost all of the text found on the internet.

2.3 While we enjoy the many benefits of using the likes of Amazon, Google and Facebook we would also like to know what is being done with the data held about us (and by whom), and ensure that our smartphones and laptops do not become surveillance devices used by companies entirely for their own profit, or by governments to control their people.

2.4 Personal data is profitable for companies who sell advertising which targets selected groups of consumers, especially now that machine learning algorithms can make highly accurate predictions about the personality, economic status, lifestyle or likely political or religious leanings of any individual, using readily available data such as internet browsing history or "likes" on Facebook. Some have labelled this "Surveillance Capitalism", a whole new market now being exploited in which we (or, more specifically, our attention) is the product. Legal protection is required for individuals against harm and discrimination in a situation where a few know everything about us while we know little about them.

2.5 While the development of AI has been an area of academic interest for many years, it is now fully within the realm of commercial interests, multi-national companies, and political parties. The influence of profit motive on the future direction of AI integration within society and our individual lives is already causing ethical dilemmas and conflicts for society. One example of this is the way algorithms are more likely to promote negative, fear-based news stories or posts to readers because they will illicit greater interaction and engagement. While this positively affects advertising income, it negatively affects social stability. Another example is the research that reveals how a major social media company's dataset provides, with analysis of a few hundred "likes", better knowledge of your personality than your spouse has of you.

2.6 The debate about regulation of AI algorithms is centred around three main issues: fairness, transparency, and accountability.

2.7 When a computer running machine learning algorithms is used to analyse medical scans or predict the shopping habits of a retail customer, the computer must first "learn" to recognise patterns from sets of training data. Machine learning is only as good as the data they are trained on, and data is blind to the human stories behind the statistics. Human biases and prejudices around gender, race, religion, lifestyle and postcode can become embedded and hidden in population data, so that when these data are the basis for generating profiles of jobseekers or loan applicants, or employed by law enforcement to screen for likely criminal activity, the results can be profoundly discriminatory and a risk to human rights. **Fairness** in Machine Learning is an active area of research which aims to develop software which can detect and eliminate such hidden biases.

2.8 To have **Transparency** in algorithm development, access must be allowed to the computer code on which a decision is based, and an adequate explanation be provided in language understandable to the general public.

2.9 **Accountability** is holding someone responsible if an AI system causes damage. Programmers should allow for human oversight able to override machine-made predictions and be held responsible for results especially decisions on credit worthiness, medical diagnoses, criminal profiling and all other areas where the risk of harm is potentially severe.

2.10 These are clearly areas of ethical concern, and over the last few years the EU and the UK have enacted relevant legislation, with further regulation proposed. However, it is essential that the churches engage in these debates, as has Pope Francis recently: "AI-based technology must never be used to exploit people in any way, especially those who are most vulnerable, Instead, it must be used to help people develop their abilities and to support the planet" (Rome Call for AI Ethics, 2020).

3. **When Computers Do Things We Normally Expect A Person To Do**

3.1 AI is being helpfully deployed in many instances to make the world healthier and more sustainable, but when machine learning algorithms attempt tasks that have up to now been considered the preserve of human beings, it raises many questions. What is special about being human and what is different about "machine intelligence" compared to "human intelligence"? Will machines outsmart humans one day? Some people foresee a future when super-intelligent machines, with cognitive abilities outstripping those of the human brain, become an existential threat to the human race.

3.2 However, it is worth considering how the "intelligence" of machines is quite unlike the way humans think and solve problems. AI algorithms create outputs which are purely mathematical in the form of predictions expressed as statistical probabilities, derived from the rules and patterns found in training data. AI is not imaginative and cannot generalise or make assumptions about a situation if the information is not already in the training data. Differences between computer algorithms and the way the human brain functions go very deep.

3.3 Compare, for example, the operation of a surveillance camera with how a human person goes about looking for someone's face in a crowd. The surveillance camera will capture a facial image and compare the metrics with other faces, perhaps derived from social media or other picture archives, and will identify the closest match. In contrast a person looking at a face in a crowd recalls memories and generates emotions that make them feel pleased surprised, angry or indifferent towards the person they see. As a social being with a sense of self and talent for empathy, the observer can share their perceptions with others employing abstract concepts using the subtlety of language. In short, they establish human to human relationships with an awareness of the uniqueness, value and needs of another person. AI has none of these skills and only imitates human thought without comprehension, lacking the consciousness necessary to create an original story.

3.4 Theories of consciousness which have been proposed include Integrated Information Theory, Global Neuronal Workspace, as well as others which argue that consciousness can never be explained by neuroscience because it is subjective. Consequently, we do not know whether a future AI machine could ever be considered sentient.

3.5 In 1950 the mathematician and pioneer of AI, Alan Turing, introduced a test of machine intelligence to decide whether or not a person can be fooled into believing that a computer is human. Based on the Imitation Game, a party game for three players, a questioner, sitting in a closed room, communicates with two other closed rooms, in one of which is a person and the other a computer. The questioner tries to determine which of the two other players is a computer. An impressive milestone in computer technology was passed when this test was first claimed to have been fulfilled in 2014.

3.6 But if the future is to combine machine intelligence with human creativity, how do we relate to a computer running an AI algorithm, perhaps embedded in a mechanical robot, that becomes part of our daily life as an assistant, companion or even therapist? What roles may computers have in the life of the Church, for instance in pastoral care? Perhaps surprisingly, we find that AI algorithms can be successfully employed to do some tasks we might think of as pastoral or relational and only able to be done by a person: for example talking therapies for anxiety and depression.

4. Chatbots And Online Therapy

4.1 Depression and anxiety are major public health issues worldwide and low-cost, easily accessible forms of treatment are urgently required. Both drug treatments and talking therapies are effective treatments for depression and it seems that both types of treatment act in similar ways to bring about changes which lead to recovery from depressed moods. Internet-based cognitive behaviour therapy, even without the direct involvement of a human therapist, has an important place in treatment.

4.2 Covid-19 has arguably accelerated the phenomenon of computerised psychological therapies, in which there has been a growing interest in recent years. One major change in 'digital therapy' has been two humans using a computer to communicate in the context of a talking therapy. In contrast, here we are going to focus on the situation where instead of a human therapist, there is a computer at the other end of the relationship.

4.3 As far back as the 1960s, computers were first predicted to take on the work of psychologists. ELIZA was programmed so that whatever the client says to the computer is met with supportive comments inviting further comment ("Tell me more…") or mirroring ("So you say that you're feeling a bit low today…"). While this was amusing for a while, it swiftly became clear that there is no intelligent being at the other end of the relationship, or at least only a rather frustrating and inflexible one. This was an early example of a 'chatbot,' and many commercial companies now offer a 'chatbot' facility on their websites for customer services as an alternative to speaking directly to a human adviser. Recently in Taiwan, AI has been used to support people in quarantine because of Covid-19.

4.4 Of all talking therapies, cognitive behavioural therapies (CBT) have been the most computerised of the talking therapies and the UK National Institute for Health and Care Excellence, in 2020, recommended greater access to digital therapies for people with depression. CBT is based around the recognition that our thoughts, feelings, and actions are all related to each other and, most commonly, that if one deliberately tries to think or act differently, it can make one feel better. The structure of CBT lends itself to the therapy being transferred to a computer and current evidence suggests that computerised CBT is as effective as similar therapy conducted face-to-face between a therapist and client. Consequently, some very sophisticated internet-based therapies have been developed. For example, Wysa is "an AI-based emotionally intelligent mobile chatbot app aimed at

building mental resilience and promoting mental well-being using a text-based conversational interface."

4.5 Another approach to the application of AI to mental health involves social media. There have been calls for the use of the masses of data many people provide to social media companies to be used to further mental health care and research, including the potential to identify people who seem to be low in mood and potentially suicidal based on the patterns of what they post and what they write on social media platforms.

4.6 While there are some important ethical concerns to be addressed about privacy and sharing sensitive healthcare data with commercial companies, we have surprisingly strong evidence that chatbots based on AI algorithms can be as beneficial to clients as person to person encounters in the treatment of depression using behavioural approaches. AI technology has the potential to provide effective, affordable, accessible treatment for sufferers of serious depression and anxiety, across society from low-income countries with poor health infrastructures to places with well-funded health care systems.

4.7 While robots and chatbots lack inherent spirituality – something that we find in person-to-person contacts or in the relationship between a person and God – and will not replace carers or therapists, they may yet have a beneficial place in helping people to explore their own spirituality in situations where interpersonal contact is difficult. This may be the case in bereavement counselling in areas where no individuals have the necessary training, for instance. Thus there are real opportunities for the Church boldly to explore the use of AI technologies in areas such as pastoral work, counselling, worship, and mission.

5. Autonomous Cars

5.1 Fully autonomous vehicles are not yet permitted in the UK, but many cars already possess some level of autonomous driving. The systems being developed are based on AI and machine learning algorithms. For many people, who might be required to share the motorway with or ride as passengers in self driving vehicles, this new technology may seem unfamiliar, strange and threatening, raising challenging questions about safety and accountability. For example, who is responsible when a car, controlled by driver assistance, is involved in an accident? Is it the person in the driving seat, the person who programmed the AI algorithm or the car manufacturer? Who takes the moral decisions in a crisis situation, "choosing" for instance, whether to take action which protects the car occupants instead of avoiding pedestrians who may be hit by the car as it swerves away from another impact?

5.2 The first step in automated driving is likely to be the introduction of an Automated Lane Keeping System, which steers a vehicle in lane. The Law Commission and the Scottish Law Commission are currently consulting on a regulatory framework for automated vehicles. Among their proposals is the idea that while a vehicle is driving itself, the person in the front seat is not considered the driver and will not be prosecuted following an accident: instead, responsibility will be with the manufacturer or operator who registers the vehicle and ensures its safety on the road. When the human in the driving seat takes control of the vehicle they will be responsible in the normal way and there will be a clear transition procedure for handing over to or taking back control from the automated driving system. In these discussions the public are being consulted and ultimately they must be convinced that automated vehicles bring clear benefits at an acceptable level of risk before this new technology is fully accepted.

6. Killer Robots

6.1 Fully autonomous weapons which can select and engage targets, completely without human intervention, do not yet exist, but semi-autonomous weapons, operated by computers running AI

algorithms with little human oversight, have been deployed in some countries. However, we cannot doubt that countries around the world will be developing fully autonomous weapons systems at present. A particular concern is the speed at which decisions are made, particularly when the data upon which those decisions are formulated is not complete, as is usually the case in conflict situations. Having a human component, while it may remove the advantage of speed in decision making, may help to avoid mistakes that would have a costly toll.

6.2 Campaigners against their use, including many faith groups, question the morality of these weapons, claiming they break international humanitarian laws relating to respect for human life. Among member states in the United Nations there is widespread support for a ban on killer robots and with the support of more than 30 countries, the **Secretary General, António Guterres**, has proposed a new international treaty that would pre-emptively prohibit the development and eventual deployment of these weapons, stating: **"Machines that have the power and discretion to kill without human intervention are politically unacceptable and morally despicable."** Talks about regulation continue at the United Nations but with little sign of progress at present. However, the relevance of the doctrine of fairness, transparency, and accountability in this area is clear.

7. Made In God's Image

7.1 The Church has, should it choose to accept the challenge, a particular insight which she can offer to the world as we, globally and locally, incorporate AI into our activities, whether seen or unseen. We have millennia of reflection and thinking about what it is to be human, the ramifications of hubris and idol worship, and the effects of elite-driven change upon the poor and marginalised.

7.2 But a much larger question which AI poses, particularly in forms that are more overtly relationship-based, is the simple and Biblical question: "Who am I?" It is a question asked of God, of John the Baptist, and of Jesus. It is the question of our identity and of how that identity is formed, given meaning and purpose. This is a question which reaches beyond corporeality, or binary digits, to a spirituality framed within a creative and abundant universe: one in which relationships amongst creatures themselves and with the divine presence is the foundation for right living. It is the question that hangs in the air in the poetic opening chapters of Genesis as they explore the place of humanity within God's creation and what it means to be made in the image of God.

7.3 Along with this question of identity is our understanding of consciousness, of freedom, of the ability to make decisions and be held accountable for the decisions we make, lest we be ejected from our own Eden. This draws into the discussion theological concepts such as self-sacrifice and grace, forgiveness and thankfulness. All of which are predicated on living in relationship with each other, of a sense of each other's worth, and of the inalienable rights that we each possess by nature of our being.

8. Consciousness Or Mimicry?

8.1 At the moment there is much debate about whether it will be possible for a future AI to become so complex that a form of consciousness might emerge or whether an AI will only be mimicking consciousness, albeit well enough that it may fool most humans, most of the time.

8.2 There is scope for both good and evil, either way. An AI which develops consciousness but is constrained only to serving the ends of others with no choice about these ultimate goals is trapped in a form of slavery. An AI that develops consciousness and takes an interest in sci-fi movies such as 2001: A Space Odyssey, Blade Runner, The Terminator, or The Matrix, might formulate its own priorities and purposes beyond those which humanity would wish to impose upon it. What can be of little doubt is that an AI that develops its own sense of identity, within a Christian theology of creation, becomes a new creation within God's creation, whether artificially, naturally or

supernaturally formed. With that new identity will come rights and responsibilities both on the part of a sentient machine and on humanity.

8.3 An AI, on the other hand, which is able to mimic relationships could and would be used by governments, companies, and movements (not least the church) in order to manipulate and influence people. One might argue that this is no more than an outworking of the natural desire to change people's minds: a mission orientation. However, in the particular instance of the use of AI it could lead to further entrenchment of the haves over the have-nots, of the wealthy elite who control these AI becoming more and more powerful. We have already witnessed the beginnings of this with the use of massive datasets from social media sources and their accompanying advertising algorithms to influence people over social and political issues. It is notable that rich, white men are the driving influencers for much AI development.

8.4 Tools are now available that provide churches with ways to utilise these same strategies to target particular groups for their own purposes, in some cases targeting vulnerable people with the purpose of indoctrination into political extremes.

8.5 As with the issue of autonomous weapons, an AI that is able to mimic human responses will not suffer or feel pain, nor know the personal cost in terms of relational damage and shame, of actions that it may cause which do harm to others. It is possible to build reward feedback loops which may help it to learn and adapt its behaviour to alleviate these risks, but that is not the same as a human being's sense of shame or feeling of remorse.

8.6 In the same way, actions which do good such as self-sacrifice are essentially meaningless to an AI that may be readily recreated from a backup of its processes and algorithms. One might, cheekily, point to Jesus' sacrifice on the cross as just such a situation, in which he knew his life was safe in the hands of God. A big difference is that Jesus chose to take the path to the cross, knowing the pain and anguish it would cause him and those who would bear witness to it, in the hope and trust that it would be, ultimately, worth it for the greater goal of revealing God's steadfast love.

8.7 It is fascinating to note the parallels to Christ in the portrayal of the replicant Roy Blatty, an advanced AI in the film Blade Runner, who approaches his moment of termination with a poetic soliloquy on his experience, having just saved the life of the person who was seeking to terminate him, lamenting that his experiences are like tears in the rain. The director, Ridley Scott, is known for his exploration of religious themes in his films.

9. Transhumanism

9.1 Can the human brain's performance be enhanced in some way by directly linking it to a computer? If successful, the development of brain-machine interfaces, a project backed by Elon Musk through "Neuralink" could introduce effective treatments for some intractable neurological conditions. Threads of miniature electrodes are surgically inserted into the brain and linked to an AI machine, able to record and analyse underlying brain electrical patterns while at the same time being able to deliver activating stimuli to specific brain structures with great precision. The technology has enormous potential for controlling prosthetic arms and legs, treating degenerative diseases and restoring function after trauma, for instance restoring a lost sense of touch or even vision and restoring to humans the ability to walk again after spinal cord injury.

9.2 The desire to see such development is unquestionably well intentioned, but the potential for misuse is also apparent. Some extend the ambitions of brain-machine interface technology and aim for a utopia when the limitations of ageing and even death are avoided once and for all. This is the area of transhumanism: the transition from carbon-based life embodied in our fleshy bags of bone, muscle, and neuron, into a silicon (or whatever substrate forms future AI) based existence. Stretching the metaphor of the brain as a computer, transhumanism proposes that a person's memories and thought patterns could be "downloaded" to a super-intelligent machine enabling that

person to exist beyond the limitations of their bodies, with extraordinary potential to change how they would then relate to other people and the environment.

9.3 Again, this raises the fundamental question of what it is to be a human being. If all we are is an intricate web of electrical activity in the neurons of the brain, then downloading that pattern of interconnections onto a computer could lead to extreme longevity, if not quite immortality in cyberspace. Reductive materialism offers hope for near-eternal life in the presence of a super-intelligent machine. For those who have witnessed a loved one's brain being affected by dementia, the appeal of being able to overcome the degradation that will follow is immense.

9.4 Even if technology may in the future allow us to transfer memories and thoughts to another substrate (a highly unlikely scenario given how embodied our brains are), there is no guarantee that what is transferred to a computer would be human as we know it. Would it still have a soul? Perhaps, but would it still be a human soul, or should it be considered an entirely different species? Indeed, this movement towards transhumanism in part to avoid death perhaps tells us more about the nature of humanity, our frailties, and concerns. A challenge to transhumanism is that it is a form of idolatry, that it fails to understand that part of what it is to be human is to grow through times of struggle, pain and sacrifice. It is also extremely individualistic, seeing value only in the worth of self. In the desire to 'live forever' it is easy to lose the importance of living in community where we pass knowledge from generation to generation, struggle on behalf of future generations, and see wealth as something that can foster growth for others. This is another area in which the Church can offer a helpful and positive alternative vision to what might become a soul-less, joy-less eternal transhumanist existence.

10. Concluding Theological Response

10.1 There is no doubt that AI is here to stay and it may develop in ways which we can only hesitantly sketch, possibly towards some form of sentience. It will bring great benefit, but, as with nuclear power, there may be huge costs or potentially destructive forms of AI with which humanity will need to grapple. We already see that with the effects of AI on financial markets, on the development of surveillance capitalism, where we are each commoditised. For the Church, the place of AI within a bigger narrative of redemption and reconciliation for Creation and relationships, and of its role in the building of the Kingdom of God, should be of concern not just to theologians, but to ministers and congregations.

10.2 While parts of the church have practised theologies of supremacy, a hubristic 'our way is the only way', and escapism, 'get to heaven as soon as you can to escape the horror of earth', these limited theologies are inadequate responses to the development of AI within society and Creation.

10.3 An alternative theology is one which is participative. This is not a theology of dominance over Creation, nor is it a theology that seeks to abandon the planet to whatever existential crisis is popular for that generation. Rather it is a theology that seeks engagement in the world and to find ways to bring healing and reconciliation, to build up the Kingdom of God on Earth through our actions, using Jesus' teaching and life as our guide and inspiration. This alternative is community-based, not domination-based, and encapsulates the transformative power of people working together under the guidance of God. It means engaging honestly with difficult questions, being humble about what we do not know, and seeking to learn from each other. Of course, this is the kind of model that Jesus himself used, so it has good form.

10.4 When we think about ourselves as made in the image of God, it is only natural that we tend, in a Trinitarian model, to look to Jesus as our exemplar. What the development of AI may demand of us is that we take more cognizance of what it might mean to be made in the image of the Creator, the one who brings into being new forms of life from the dust (and silicon) of the earth and how that shapes our responsibility to AI and those who will be affected by AI. Furthermore, seeing ourselves

in the image of the Spirit of God, an altogether different form of existence from that which we experience, may give us insight into what life in a digital age might encompass.

10.5 Seeing our interactions with technology and our involvement in the development of AI through a relational lens such as that of the Trinity may help us to engage productively with these developments, so that we can help shape AI that God would declare to be good.

10.6 The world and Creation are not static. Developments such as AI will always arise as humanity deploys its inherent creativity. A sign of the Church's maturity is how it responds, engages with and tries to assist communities in how they move forward with AI as part of their lives.