Trusting artificial intelligence for good: How at Rolls-Royce we are stepping beyond theory with practical, applied ethics for industry and beyond

At Rolls-Royce, we are probably best known for our advanced manufacturing and jet engines rather than our developmental work in Artificial Intelligence. But we have more than 30 years of experience in advanced data analytics and have been developing our AI capabilities for many years.

Later this year, we will make freely available our work on how to construct an ethical framework for the use of AI, and how that framework can be used to generate AI in which people can genuinely have trust.

I believe the impact of our work has the potential to go far beyond just our own field of expertise into broader industrial AI applications and well beyond industrial application.

This is a complex area and, of course we haven't solved all the challenges presented by AI.

But during our peer review process with experts from big tech companies, academia, automotive and pharmaceutical industries, we were genuinely surprised to learn that our work – which began as a way to address an internal challenge – stands at the leading edge of the practical and ethical application of AI in a critical industrial context.

And as we went on, we came to recognise that its implications could extend way beyond that.

Why am I talking about something like this? These days my job is to be the CEO, but I began as an engineer and I've worked in the world of technology since.

AI was an exciting new world when I graduated, but it stumbled then as it had done before, due to a lack of computing power. I've been privileged to be part of the semiconductor world through the vital decades which have delivered the progress to fix that and enable AI's subsequent resurgence.

As with all technology there is a lot of hype. It's true not all applications of AI are useful. Today, for business, and for the good of society we need to be thoughtful about how we apply it. Then like most of the other tools we've picked up since humans have been humans it will indeed be useful.

Certainly, in wider society AI can scare people, and there are scaremongers out there. Like most of the other tools we use, if misused of course there is a dark side. But I believe we can do better than that, dismissing a new technology or tool because of the scary bits won't help, and isn't justified.

Instead focus on the good bits and it can enable real scale; it's the same as when you grow a business by recruiting and delegating. Then we unlock human potential; and accelerate greater industrial, commercial and societal endeavours.

Of course we need to reassure about those good bits vs bad bits and our communities and colleagues must be able to really trust artificial intelligence and believe that it's being applied for good outcomes; and this is at the heart of the work we have done at Rolls-Royce.

Let's look at the Roll-Royce experience.

Our industrial journey in Machine Learning, Artificial Intelligence and Advanced Analytics started in 1980 with engine test bed measurement and analytics. It really came to life as a disruptive technology for us in the late 1990s with our Engine Health Monitoring system. That has been a key enabler for the delivery of many billions of pounds worth of aircraft maintenance contracts.

To give you a sense of scale of what we do: we can monitor 6000 to 8000 flights every day, which equates to more than 3000 Rolls-Royce engines in the sky at any one time. Each engine has multiple sensors on board that relay information back to us during the flight.

In total, we analyse more than 5 million data parameters from our engines every day and use AI to provide insights to our engineers. If we spot anything out of the ordinary, we can dispatch a maintenance crew to be ready on the ground when the aircraft lands.

So what does the AI that we use actually do? It learns the normal operational state of our engines and scans for relative system changes and looks for unknown anomalies, rather than relying only on a preconfigured set of fault signatures. This means our engineers can make decisions quicker than ever, reducing disruption for our customers.

But it's not just about aviation and the behaviour of our engines. Our current work includes applying AI, through our in-house data innovation team, R2 Data Labs, to improve the management of risk in our supply chains; to predict market demand signals; and to examine how to improve the efficiency of our operations. We've deployed this for example with the microgrids created by our Power Systems business, making our industrial power technology more reliable and sustainable.

In the future, our AI work will involve increasing our use of cloud-based services which will be governed by a data ethics framework that covers more than 200 of our digital projects. And we will continue to rely on AI to efficiently design and operate our products.

Due to the markets which we serve, the key throughout all this work is safety. About 60% of our revenue is about Aerospace where that is obvious, but building on that DNA and expertise we differentiate in the other the market sectors in which we operate.

So we are a company where having a safety mindset is fundamental to our products and services, and so it must be fundamental to our data innovation. I believe that is why we're the first to go beyond theory in this space, so that we can actually demonstrate that the AI systems we use are trustworthy and safe.

I used to work for ARM, the company behind ARM microprocessors. Most people have never heard of ARM, but there are hundreds of ARM processors in all of our lives every day, in most cases we just don't know they are there.

When you get on an aeroplane and it takes off you can feel the power of the engines, and that's what we associate woth the engine. But similar to the invisible ARM processors you are just not aware of all the technology in that engine, and elsewhere on the plane that is keeping you safe.

Our AI is deeply embedded into other people's products and services, and it's BA or Lufthansa that fronts up to the consumer. Rolls-Royce's AI doesn't feature in a consumers' understanding of how the digital world is changing their lives.

We need people to trust it, as they trust the power of that engine to propel the plane into the air, but they don't even know the AI is there.

The current debate about the use of AI is very much focused on the consumer and the treatment of consumer data. But we believe that what we have created – by dealing with a challenge rooted squarely in the industrial application of AI – will help not only with the application of AI in other industries but far more widely.

Let me explain. Over the last year we have been solving an internal business critical challenge involving demonstrating the efficacy of robot inspections. In reaching a solution, we have achieved two breakthroughs. And, while our aspiration is to become the world's leading industrial technology company, the impact of our work extends to all uses of AI.

Firstly, we have developed a framework that ensures the ethical application of AI – one of its biggest public relations hurdles. This is a practical way for any application of AI to be ethical so that it can be used to contribute fully to the growth, wealth and health of our world.

Secondly, we have gone beyond theories and concepts to build and implement a step-by-step process that is robust enough for us to completely trust our AI and its outputs – whether for an ethical process such as sifting CVs or a critical /safety critical business process such as robotic inspection of components. It can be implemented immediately by focusing only on the outputs of algorithms, not the algorithms themselves, which are constantly learning and evolving.

My colleague Lee Glazier, who has done a lot of the development work, will be discussing both these aspects in more detail in his presentation at noon today and I invite you to watch it.

From an industrial point of view, we know this is only part of the solution to how to apply AI with trust and within an ethical framework, but it is a big step in moving us and the industrial world towards industry 5.0, where artificial intelligence is applied more broadly and creates a shift in the contribution that humans make to industrial processes.

We need to be Tackling key issues up front.

So what does this industry 5.0 mean for workers around the world and being selfish specifically for us at Rolls-Royce?

Its potential impact on people and particularly on loss of jobs, is one of the most emotive aspects of the AI discussion. As part of our stakeholder engagement, we have taken our union representatives in the UK through the work and received a positive response to how we are tackling this issue.

There are also existing research papers, such as by the World Economic Forum, indicating that by 2022, nearly twice as many jobs will be created than those lost, by AI-enabled automation.

And what about regulation? The current approach is moving rapidly towards what has been described as 'opening the black box' to check the algorithms themselves. In practice this requires a business to reveal what is often sensitive and proprietary work. It's also quite a hard thing to do for a regulator and perhaps most importantly it only provides a snapshot of what an algorithm that is learning, developing and adapting to its task, is doing at one particular point in time.

Our simple, five-step process focuses on validating the outputs of algorithms so there is no need to crack open the black box. It can be applied now, at comparatively little cost, and with great efficacy because the checks can be repeated every few minutes as the algorithms evolves.

We are not seeking a commercial return from this work beyond that which its use will achieve for our business and customers. We have come to realise that it's potentially so useful that everyone needs free access to it and as a CEO I'm a strong believer in open source sharing, open architectures, integration and networking.

I realise that some things you just need to give away and trust in the business benefit materialising through overall growth.

As a business we are open to collaborative innovation and we will continue to talk to key stakeholders, customers, counterparts and technology leaders to share our work in detail to see how we can help each other progress for the greater growth, wealth and health of our world.

For example, our specialists are already collaborating with the National Hospital for Neurology and Neurosurgery in London, researching potential neuroscience techniques that could be applied to identifying any mutations or unintended biases in AI systems.

Lee Glazier will also be referring to this important work at noon, including our research into parallels with genetic mutations.

There is much more to do and by publishing our findings we want to move the AI ethics conversation forwards from discussing concepts and guidelines, to accelerating the process of applying it ethically.

There is no practical reason why trust in AI cannot be created now. And it's only with the acceptance and permission of our society – based on that trust – that the full benefits of AI can be realised, and it can take its place as a partner in our lives and work.