



Australian Government
Chief Scientist

DR ALAN FINKEL AO

**Institute of Public Administration Australia (IPAA) and
Australian Council of Learned Academies (ACOLA)**

**The role of AI in Australia 2030:
policy implications and the nature of industry in society**

AI: Can we get it right, please?

**Tuesday 18 August 2020
Online keynote address**

It seems fitting to open today's discussion by talking about the second most powerful computer ever built.

Designed by hyper-intelligent pan-dimensional beings, and programmed with an enormous dataset – literally every piece of knowledge in the known universe.

Built for one purpose: to discover the answer to the ultimate question of Life, the Universe, and Everything.

After seven and a half million years of non-stop processing, the great supercomputer, aptly named Deep Thought, stirred and, with infinite majesty and calm, announced that the long-awaited answer was...

42.

When pressed further about the meaning of the Answer, Deep Thought responded obliquely.

The great supercomputer admonished the hyper-intelligent pan-dimensional beings, noting that the answer given was unquestionably correct, but perhaps the focus should not be on the answer but on the question that they asked.

Now, it cannot be a coincidence that I share Deep Thought's sagacious answer with you exactly 42 years since Douglas Adams' classic story, 'The Hitchhiker's Guide to the Galaxy,' premiered as a radio series on BBC.¹

Like the hyper-intelligent pan-dimensional beings who programmed Deep Thought, I trust that human AI researchers – while perhaps not being quite as smart or multidimensional – will have a crystal clear vision of their ultimate goal: to create AI that is capable of reasoning, planning, solving problems, thinking abstractly, comprehending complex ideas, learning quickly and building on experience.

In short, to produce human intelligence without the blood, tissue, and goo.

But in our eagerness to reach this goal, we often neglect to ask the right questions.

And as Deep Thought reminds us your answers are only as good as your questions.

Determining the questions can be difficult.

Don't panic.

Throughout my term as Australia's Chief Scientist, I have been privileged to learn from expert stakeholders.

And now, with this likely being my last major speech on Artificial Intelligence, in my current role, I hereby declare that it is altogether fitting that I recap the key questions that have framed some of my thinking on this issue.

Questions that will need to be kept front of mind to ensure our ingenuity is forever rooted in our values.

Last year, I had the opportunity to launch the artificial intelligence horizon scanning report from today's host, ACOLA, an organisation that brings together the hyper-intelligent *single*-dimensional intelligence of the Fellows in our Learned Academies.

It's a fantastic report, the best of its kind that I've seen.

And it serves us as a foundation for the future.

The report notes that there's no magic solution to any policy challenge – least of all the innovative applications of artificial intelligence.

Instead, it all comes back to the essential foundation of trust.

Are we capable of extending our trust to AI?

To invent the modern world, our ancestors had to devise the complex web of laws, regulations, industry practices and societal norms that make it possible to trust our fellow humans.

So what will it take for you to trust artificial intelligence? To allow it to drive your car? To monitor your child? To analyse your brain scan and direct surgical instruments to extract your tumour? To spy on a crowd and zero in on the perpetrator of a robbery?

At a Creative Innovation conference in 2017, I asked the audience to think of the sophisticated spectrum of rules that operate in human societies.

At one end of the spectrum, we have accepted societal norms – manners, if you like – and incentives for good behaviour. Please and thank you, rewarded with a lolly.

Moving along the spectrum, we have organisational rules that govern how to act in classrooms or the workplace, penalties for misdemeanours such as parking in a Loading Zone, then moving further along custodial punishments covering crimes such as robbery and assault.

Approaching the other end of the spectrum, we have severe punishments for the worst civil crimes such as premeditated murder or terrorism, and finally, there are internationally agreed conventions against weapons of mass destruction.

We will need a similar control spectrum for AI.

Some applications of AI, such as medical devices and car manufacturing, will necessarily be covered by existing regulatory regimes.

But at the societal norms end of the spectrum, there is a worrying void. After all, what constitutes “good behaviour” in a social-media company’s use of AI?

It’s amazing how much information we’re willing to give up in exchange for a bit of convenience. Information that can and will be used to target us with advertisements and coercions.

We have little protection.

End user agreements exist to protect the vendor, not the consumer. None of us has the knowledge or the time to individually vet the AI products we use.

When was the last time you read the terms and conditions before clicking “Accept”?

In our human intelligence society, there are consequences for falling short of standards.

We need the same for AI developers. A way for consumers to recognise and reward ethical conduct.

The most straightforward method would be an AI trust mark, allowing consumers and governments to identify and reward ethical AI; similar to the globally recognised “Fairtrade” label that today adorns products such as coffee, chocolate and tea.

Let’s call the AI trust mark the Turing Certificate, after the great Alan Turing.

A Turing Certificate would indicate that the vendor and its product are worthy of your trust.

I have recommended the Turing Certificate in several speeches, and in discussions with Standards Australia. Some tentative steps in the right direction have been taken.

Is this a global measure that Australia could help to foster?

I argue that we should.

We have more to gain than most.

Where we compete in the global market, we compete on quality.

It’s true in agriculture: our reputation secures a premium price.

It’s true in higher education: our reputation is the foundation of our third biggest export industry.

And it should be true in technology.

A system that rewards quality and prioritises ethics will reward Australia.

But I emphasise: this voluntary system would only be suitable for low-risk consumer technologies.

What about AI applications that have direct impact on our freedom and our safety?

Could Australia be influential in developing ethical frameworks in these more significant applications of AI?

In that light, we are already seeing a cohort of leaders across government squaring up to their responsibilities as AI adopters and custodians of human rights.

For example, last year, the Australian Government released a draft set of principles that, if implemented, would reduce the risk of negative impacts from AI products and encourage best ethical practices during development and use.

Not to be outdone, the Human Rights Commission, under the leadership of Ed Santow, is deep diving into the difficult issue of human rights and digital technology. It is a three year project and I am proud to have been on the advisory committee from the beginning.

The pithiest encapsulation of human rights, in a world where AI exerts formal power over us, was expressed by Michael Pezzullo, the Secretary of the Department of Home Affairs. He proposed a line in the sand, not just for border security but for every decision made in government that touches on a person's fundamental rights.

He called it "the Golden Rule".

"No robot or artificial intelligence system should ever take away someone's right, privilege or entitlement in a way that can't ultimately be linked back to an accountable human decision-maker."

To me, this Golden Rule, which I have quoted in many speeches, is the mark of a public sector fit to be an ethical custodian.

It underscores that as human beings, we should always have the right to appeal decisions made by artificial intelligence to a human adjudicator.

We must all be involved in the national discussion about how we integrate artificial intelligence into our society.

I noted in my keynote speech at the launch of the Human Rights Council international conference in 2018, that every time we come to a decision point we must ask ourselves the fundamental question:

What kind of society do we want to be?

Nations like Australia have choices.

We all bear great responsibility to safeguard our time-honoured rights and values.

In my mind, a fair-minded people can build artificial intelligence into a better way of living, without needlessly encouraging a conflict between science and ethics.

But how?

We must start with good legislation, guidelines and ethical behaviour.

But we can go further by using yet more technology to mitigate the problems caused by the current generation of technology.

Let me give you an example.

In a speech earlier this year to launch the University of Melbourne's Centre for Artificial Intelligence and Digital Ethics, I noted that, currently, nearly all the information about what we are doing on AI platforms is stored, deconstructed and analysed by cloud-based servers.

These cloud based servers are devoid of any morals.

They are servers not servants.

As such, they present an ethical dilemma, which is that I want the immense benefits that they provide; but I am alarmed that in order to do so, my smart device relies on server-based AI in the cloud. AI owned and operated by corporations whose *own* interests far transcend their commitment to *my* interests.

From the cloud, these corporations can identify me, follow me around, send me advertisements, and potentially share my information with third-party organisations.

From the daily news, it would seem that appealing to these corporations' better judgement, passing legislation and issuing guidelines does little to shift their business models.

But, I see a future in which *more* technology will give me what I want. An AI server with virtually unlimited power that is as loyal to me and my interests as my ever loyal pet dog, Bessie.

In the long-term, I expect that the software and the processor in my phone will each be a thousand times more powerful, so that my phone will take my questions and interpret them locally.

Then, and only then, will it reach out to the cloud to get the answers I need.

It will reach out anonymously.

So that my privacy need never be compromised.

AI on my device, not in the cloud.

I urge you to reflect on these questions, with – dare I say it – deep thought.

Hyper-intelligent pan-dimensional beings may have discovered the answer to Life, the Universe, and Everything, but the ultimate question we humans must ask of ourselves is the one I have frequently posed:

What kind of society do we want to be?

Answering this question might not take seven and a half million years but it will take time, and constant vigilance.

We must all endeavour, with determination, to harness the power of scientific progress for the benefit of our society, while safeguarding the ideals of our society.

So my challenge to this gathering is simple: Show us your path to the ultimate goal: *to define the kind of society we want to be.*

Inspire us, as a nation, to achieve it together.

I wish you all the best for this vitally important discussion.

And with my term as Australia's Chief Scientist coming to an end this year it is altogether appropriate for me to say:

“So long, and thanks for all the fish.”

May the Force be with you.

Thank you.

ⁱ <https://www.bbc.com/historyofthebbc/anniversaries/march/hitch-hikers-guide-to-the-galaxy>