



**Australian Government**

**Chief Scientist**

**DR ALAN FINKEL AO**

**Can 'Natural Law' underpin 'Artificial Intelligence'?**

**Speech to**

**Joint Conference of  
Federal Court of Australia and Law Council of Australia**

**Thursday 28 November 2019  
Melbourne**

*\*Dear Reader. Instances of legal terminology used throughout (to lighten the mood at a formal dinner setting) have been double underlined for clarity.*

\*\*\*

It is an incredible honour to be invited to make a presentation to the Judges of the Federal Court of Australia.

This, of course, is the only way that I would ever want to appear before you.

As an engineering student, I envied my law student friends for their opportunity to participate in the theatre of the moot court.

So now, in this hallowed location, but in the comfort of a genial dinner, with your indulgence I will take the opportunity to make up for what I never had as an engineering student.

In that spirit, may it please the Court, this is the matter of the Law vs Science, and I stand before you, your Honours, as Australia's Chief Scientist and Chief Counsel to the defendant.

As a scientist, my world depends on laws.

We have Newton's Laws of Motion, Kepler's Laws of Planetary Trajectories, Mendel's Law of Inheritance, the Laws of Thermodynamics, and many more.

All scientific knowledge rests on these *axiomatic* laws, which haven't changed in *13.7 billion years*, since the Big Bang, the very origin of our universe.

So I know where *my* laws emanate from, your Honours.

In thinking about this opening statement, however, I started pondering on the other kind of laws, *your* kind of laws, and why, for centuries, scientists have called the measurable, predictable regularities found in nature 'laws'.

The concept first appeared as a metaphor in Latin poetry, before gaining a firm theoretical presence in Ancient Rome.

Roman jurists and philosophers argued that the very essence of law rested not upon the arbitrary will of a ruler, but upon what they labelled the *lex naturae*, the natural condition that exists both in our world and in our being.

The laws of morality, like the laws of science, were therefore not derived from humanity but from nature, and stood eternal, universal, and self-evident.

Hence, to this day, the term 'natural law' is interchangeably used to describe immutable moral principles and natural phenomena.

It is a foundation that shapes both our professions; our 'modus operandi' if you will.

But, your Honours, let me call your attention to the crux of this case:

Can our old values guide the adoption of new technologies?

In particular, can 'natural law' order the behaviour of artificial intelligence in our society?

Artificial Intelligence has moved forward at such a dizzying pace that it is pushing us towards bold new frontiers of imagination and innovation.

I make this statement not based on theory or ideology, but beyond reasonable doubt.

Here's the most important piece of evidence: The integration of Artificial Intelligence has made life better for countless individuals. Across all facets of society, we're experiencing a transformation that echoes the great industrial revolutions of history.

I couldn't put it any higher than that, your Honours.

Consider the use of A.I. for A.I.

By that I mean Artificial Intelligence for Artificial Insemination, otherwise known as In Vitro Fertilisation.

In a standard IVF procedure, clinics wait for the newly fertilised eggs to develop over five days into embryos, before the doctors decide which of the batch to implant into a hopeful mother.

They judge, by its appearance, which embryo gives the best shot at a successful pregnancy.

But human doctors can't watch the developing embryos constantly – 24/7 – for five days straight.

A.I. can.

And human doctors can't be trained on thousands and thousands of hours of time-lapse footage of embryo development.

A.I. can.

A.I. is helping make that choice more reliably right now through Australian company Presagen and its pioneering product Life Whisperer.

It's a breakthrough use of Artificial Intelligence that, just this month, commenced its first clinical trials.

Or, the A.I. from Google's DeepMind division, which is now able to predict whether a patient has potentially fatal kidney injuries 48 hours before symptoms can be identified by doctors.

In *your* profession, we are seeing A.I. helping lawyers perform due diligence by searching, highlighting, and extracting relevant content for analysis; by providing additional insights and extracting key data points through analytics; and by streamlining work processes.

Fully automated creation, negotiation, execution, and filing of agreements, applications, and contracts, is saving jurists and lawyers thousands of hours of work.

It is important to recognise that A.I. is not automating the legal profession out of existence. On the contrary, A.I. is facilitating growth and productivity by increasing accuracy and optimising efficiencies.

Why should humans do repetitive or menial tasks we are ill-suited to do for long periods, when we can free up our time for higher order thinking?

If the Court pleases, let me provide an analogy.

I have a pilot's license and I can still recall, when I was training on cross-country navigation, my instructor asking me why I wasn't using the autopilot.

My reply was that I thought it was cheating, and he thundered back "No! Use all the tools at your disposal. That way you will free up your mind to deal with other higher level challenges, like watching for other planes, monitoring the weather and checking flight-critical systems."

My little effort to resist the use of technology is an example of the broader truism: resistance to the rampant march of technology is futile, and self-defeating.

Let me rephrase that. "A.I. won't replace lawyers, but lawyers who use A.I. will replace those who don't."

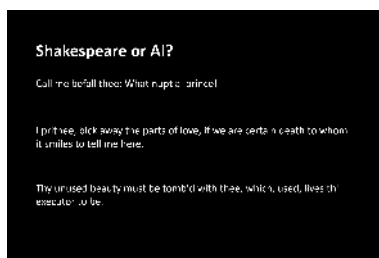
May I repeat that, your Honours. "A.I. won't replace lawyers, but lawyers who use A.I. will replace those who don't."

I say this while being deeply aware of the challenges of A.I.

At this time, I would like to direct your attention to Science's Exhibit A.

Can Artificial Intelligence match or even surpass the marvels and creativity of human intelligence?

How about we take a great human writer – let's say, Shakespeare. I'm sure we can all agree that as writers go, he was pretty good.



Can you tell if these lines were written by Shakespeare, or by A.I.?

Now let's turn to art.



Rembrandt. You've heard of him. He was a pretty decent painter.

Two of these were painted by Rembrandt. Which one was painted not by Rembrandt... but by a RemBot?

Finally, what happens when you mix easy access to increasingly sophisticated technology, a high-stakes election, and a social media giant?



These videos are expertly crafted 'deepfakes', A.I.-based technology that can produce doctored images and videos that look and sound just like the real thing.

They illustrate not only how convincing manipulated video content has become, but also the paradox that defines Artificial Intelligence today.

Like other technologies such as medicines and electricity and petrol, the same inventions that serve humanity can also cause great harm.

We cannot dismiss these deepfakes. They are powerful. They are pervasive. And they have the potential to erode and discredit the community's confidence in the integrity of A.I.

To forever shroud A.I. in suspicion and scepticism.

Indeed, I believe that only by acknowledging and confronting this reality can we ensure that the darker aspects of A.I. do not tarnish both the value and the virtue of our scientific progress.

As such, the development of A.I. provides us with an opportunity not only for intellectual growth, but for moral leadership.

That means your job and my job are fundamentally interconnected.

Through concerted and collective efforts we can fashion a framework that will enable Australia to be global leaders in the field of A.I. ethics and human rights.

Showing the world how to advance the cause of scientific research while staying true to the ideals of a prudent and virtuous society.

Just this month, the Minister for Industry, Innovation and Science, Karen Andrews, released a set of 'A.I. Ethics Principles' to build public trust, as well as help guide businesses and government to responsibly develop and use A.I. systems.

And, at the same time, the Human Rights Commission, under the leadership of Ed Santow, is deep diving into the difficult issue of human rights and digital technology. I am proud to be on the advisory committee.

As protectors of our civil liberties, your role in this endeavour will be crucial.

You can help us identify barriers to opportunity, advocate fairness and consistency in its administration, and create A.I.'s agenda for decades to come.

There is, your Honours, a factual basis for this plea.

Across Australia we still rely on a rulebook written in a different century.

As a giant of our legal history, Sir Isaac Isaacs, once stated: our laws "are made, not for the moment of their enactment but for the future...[they are] a living instrument capable of fulfilling its high purpose of accompanying and aiding the national growth and progress of the people for whom it has been made."

We therefore look to you, in humble submission, to help shape the course of this century, just as your predecessors helped shape the last.

However, there's a conundrum: we know that our laws are continually improved and adapted to reflect new conditions and realities; but we also know that technology and our reliance on its capabilities is evolving much faster than our laws.

Given this, will the Law effectively manage the behaviour of A.I. society — as it does with human society — to maximise common good and individual good?

The Law is essential to preserving order in a democracy. And we cannot have order unless people are certain of the full scope of their rights and legal protections.

As such, ambiguity over the principles that govern A.I.'s application, threatens our way of life.

As does A.I.'s vulnerability to be manipulated to give false testimony.

Think back to those deepfake videos. If, up to two or three years ago, you were presented with a single security camera's footage detailing a crime, the prima facie evidence against the defendant would have been quite damning. But today, can you be so sure?

The answer is clearly no.

But, on the other hand, this era of technological revolution has created many more sources of evidence, with phone cameras and video recorders now in abundance.

So, you will have statistical means at your disposal to determine the credibility of evidence that might not merely be tampered with, but fake from scratch.

I call your attention to another question:

Will A.I. serve to encourage and empower an ethical society or will it weaken us and drag us down into dependence and disrepute?

The A.I. we want is a product of understanding and agreement and morality, based on justice and security and individual freedoms.

But the risk of overreach — the possibility that we lose some of our core liberties in pursuit of progress — also becomes more pronounced.



Take the most common application that people think of when you ask them about A.I. in society — facial recognition.

I heard a story recently from Ted ‘Smith’, a friend. A story that emphasises just how powerful facial recognition, monitoring cameras, and connected databases are.

The tale of Ted is not hearsay. The witness was cross-examined by yours truly.

Ted was picking up his daughter at Brisbane airport, when she realised that she had had left her reading glasses on the plane.

Given his daughter had to get to a meeting in the city, Ted suggested she proceed by taxi while he would go back into the airport and find her glasses.

After checking with the airline service desk he was told to go to lost property.

As he was leaving the arrivals area, Ted saw a laptop computer, seemingly forgotten on a seat.

Not able to see an owner, he picked it up to take it with him to lost property since he was going there anyway.

A minute later, his phone rang.

“Good afternoon ‘Mr Smith’. I’m a detective from the Australian Federal Police. We have jurisdiction over this airport and are aware that you picked up an unattended laptop computer. Please proceed to the nearby Police precinct.”

Ted still has no idea how the Police knew he had picked up the laptop nor how they had identified him or obtained his phone number.

Off he went to the precinct, where he handed over the laptop to an Officer who guided him through filling in the required paperwork to hand over the laptop.

There was no threat of discipline expressed nor any mention of any follow-up.

Ted is convinced it was a test of a new facial recognition system.

This is Ted's sworn testimony. And he is an expert witness.

His story is therefore marked as evidence that the impact of A.I. in Australia is no dream of the future. It is here, now, today.

We cannot turn back the tide of technology, and we must therefore define the nature and scope of its application, or else, it will define us.

We must always remember that the same enlightened society that advanced the cause of science has also advanced the cause of justice.

The same persistence that opened up new frontiers of discovery, also opened new doors of opportunity.

As the holders of this legacy, we bear great responsibility to ensure the sacred ideals embodied in this building continue to be afforded to everyone.

And while A.I. can be a powerful aid to our cause, the consequences of any single slip-up are immense.

To focus on the basics, if A.I. is intended to make law firms more efficient, how is that consistent with the doctrine of maximising billable hours?

On a more serious note, A.I. based risk assessment tools are being increasingly used in the criminal justice system to guide sentencing.

One of the most widely used by U.S. courts is COMPAS, which has assessed more than 1 million offenders since it was first developed in 1998.

COMPAS works by analysing answers to a questionnaire that defendants must complete.

The 137 questions gather the defendant's personal history, such as whether one of their parents was ever sent to jail, or the number of their friends taking drugs illegally.

It also asks people to agree or disagree with statements such as “A hungry person has a right to steal” and “If people make me angry or lose my temper, I can be dangerous.”

Using the answers provided, COMPAS creates an assessment for risk of reoffending, which can then inform a judge’s sentencing.

But there’s a problem.

Although race is not one of the questions used, other aspects of the data may be inadvertently *correlated* to race — such as poverty, unemployment, and social disadvantage — that can lead to racial disparities in the predictions.

In fact, in 2016, the news organisation ProPublica reported extensive racial discrimination in COMPAS.

According to the report, COMPAS was deeming black defendants to be at a risk of recidivism at *double* the rate they actually were, while predicting the reverse, that is *half* of what was actually the case, for white defendants.

COMPAS' developers have since questioned ProPublica’s analysis but the fact still remains: irrespective of the amount of data, in this context, A.I. systems may reach an upper limit to what they can achieve.

With the greatest of respect, your Honours, I therefore exonerate the algorithm and convict the accomplice!

It is indisputable that human judges are biased. Studies have illustrated how legal judgments can be influenced by a range of factors including a Judge’s upbringing, decision fatigue, unconscious assumptions, the time of day, the perceived attractiveness of the individuals involved, even when and what the Judge has eaten.

Fortunately, the wide spectrum of ideologies across our legal system, spanning thousands of Judges, averages biases out.

But, as with the case of COMPAS, if a single instance of the A.I. implementation becomes so successful that it is widely adopted, then it entrenches the bias across the whole system. There is no averaging out.

To effectively combat this risk, may the record reflect that we will need to adopt two key strategies.

First, we can guard against systemic bias by ensuring that no single A.I. based risk assessment tool ever captures more than a small percentage of the market.

Second, we can go a step further by only adopting A.I. that has been methodically trained to avoid introducing biases.

At the Australian National University, Professor Genevieve Bell and her team are establishing a new branch of engineering that intends to do just that.

Making sure that the people who design and build our A.I. systems represent the myriad cultures, experiences, and perspectives that make up our vibrant society.

Aiming to deliver an authentically Australian A.I. — powered by *our* energy and creativity, and bound by *our* shared values.

But, computer engineers simply do not and cannot have the acumen needed to craft algorithms in each respective field.

I therefore move, your Honours, that you take an active and central role in the adoption and judicial oversight of A.I. in the legal profession, and broader Australian society.

Only your sound legal minds, can ensure we have a sound legal system in the future.

Your Honours, you have seen and heard Science's plea in this case.

In closing, let me review with you the key pieces of evidence presented.

First, embrace technology; resistance is futile.

Second, be at the forefront of developing solutions to the changing nature of evidence.

Third, guide the technological development and ensure that no single risk assessment tool can dominate.

So what's the verdict on A.I.?

Well that's your jurisdiction, *my* laws are axiomatic and, as such, I recuse myself from this deliberation.

But the evidence shows that *your* laws are not axiomatic and cannot be.

As Sir Isaac Isaacs rightly affirmed, the Laws of Australia are *living entities*; subjective and experiential in nature, shaped by societal decisions, and grounded in our core tenets.

And like any living entity, they evolve over time.

Our morality supports them, our righteousness sustains them, and our conviction that we are all equally entitled to inherent human rights and values, gives them vitality and force.

With that, it is altogether appropriate to say two things.

It's time to adjourn, and May the Force be with you.

Thank you.

\*\*\*\*\*

Link to report on the Boris Johnson and Jeremy Corbyn deepfake videos

<https://www.bbc.com/news/av/technology-50381728/the-fake-video-where-johnson-and-corbyn-endorse-each-other>

Link to report on Mark Zuckerberg deepfake video

<https://www.youtube.com/watch?v=NbedWhzx1rs>