Artificial Intelligence and Australia

That a select committee, to be known as the Select Committee on Adopting Artificial Intelligence (AI), be established to inquire into and report on the opportunities and impacts for Australia arising out of the uptake of AI technologies in Australia, including consideration of:

a. **recent trends and opportunities in the development and adoption of AI technologies in Australia and overseas, in particular regarding generative AI;**
b. **risks and harms arising from the adoption of AI technologies, including bias, discrimination and error;**
c. **emerging international approaches to mitigating AI risks;**
d. **opportunities to adopt AI in ways that benefit citizens, the environment and/or economic growth, for example in health and climate management;**
e. **opportunities to foster a responsible AI industry in Australia;**
f. **potential threats to democracy and trust in institutions from generative AI; and**
g. **environmental impacts of AI technologies and opportunities for limiting and mitigating impacts.**

About the Author

For the purposes of this submission Mr Rovere claims no affiliation with any organisation and makes this representation in a strictly personal capacity.

Crispin Rovere is an internationally recognised expert in strategy, alliances and nuclear policy. In 2020, Mr Rovere published the influential article: *The Artificial Intelligence Race is an Arms Race* in the United States that shaped discourse on the military applications of machine learning. Mr Rovere informed OpenAIs *Preparedness Framework* for AI risk, gaming hypotheticals for their current GPT4. Mr Rovere is the author of the brand new screenplay “Soul Code” set in a near-term future where governments debate whether to grant AI human rights.

Introduction

Artificial Intelligence (AI) is the most significant technology ever developed by humankind. AIs impact on civilisation will prove greater than the internet, combustion engine, and the harnessing of electricity. In the relative near-term, AI will completely transform how Australians live and exist in the modern world.

Until recently, AI was the exclusive purview of a small and eccentric technology elite. Since the release of OpenAI's GPT4 on 14 March 2023 AI has burst into the mainstream. Now AI dominates global discourse and is shifting market capital. Concerns such as AI safety, alignment, and disruption are now on the political agenda, including via this Committee.

This new visibility is an unequivocal good. However public discourse on the true scale, speed, and impact of the unfolding AI revolution remains rudimentary. Australia's political leaders must urgently contend with the benefits, risks, and consequences of this transformational and disruptive technology and the new industrial revolution that it spurs.

Current Disruption

GPT4 was released by Open AI in March of 2023. In just a single year the global economy has irrevocably shifted in a fundamental way, although this shift is yet to be fully felt by the average Australian citizen or policy maker.

In the short term almost all professions will be transformed, and in many cases replaced, by artificial intelligence. To warm us up, take the following narrow example:

Image generation

Today AI generates over 34 million images every day, increasing rapidly. This equates to the total number of photographs taken worldwide in the 19th century - times 100.[1] As at April 2024, the total number of AI images in circulation today now exceeds all the images ever created by humans, across all mediums, throughout all of history.

While debate rages about copyright and how AI creates work derivative from human artists, in reality this is already obsolete. AI is now innovating into new art styles and design works never before seen by humankind.

[1] Everypixel, https://journal.everypixel.com/ai-image-statistics#

AI calls this style "Vivigeometria" - a new art style that blends vivid textures and dynamic lines, merging organic and geometric elements with a bright, saturated colour palette.

Capability growth in AI generation is staggering. The AI image generator, Midjourney, released its version 1 in February 2022. Less than two years later, in December 2023, Midjourney released version 6.

Below are samples comparing the two versions:

Version 1 - Midjourney - February 2022

Portrait

Landscape

Food

Text

Illustration

Version 6 - Midjourney - December 2023

Portrait

Landscape

Food

Text

Illustration

This improvement shows no signs of slowing. In February 2024, OpenAI showcased Sora, an AI that can generate cinema-quality video of up to 60 seconds in length entirely from a text prompt.

AI generated images are in the process of eradicating the graphic design industry globally, but this is just the beginning. On a reasonably short timescale, entire feature-length cinema blockbusters will be generated entirely through AI. Individuals will produce their own hit television series on their personal computers.

Actors, directors, producers, set designers and the whole panoply of industry that currently supports visual media will, in its current form, no longer exist.

For ordinary Australians, the AI-revolution promises a creativity renaissance where individuals can produce and share entertainment previously accessible only to major studios. This democratisation means that highly motivated and creative entrepreneurs will market whole libraries of AI-generated entertainment on a scale akin to major streaming platforms such as Netflix and Amazon.

Conversely, for those who depend on specialised creative skills for their livelihood, the future is bleak. Even as AI increases their individual productivity, demand for their services is falling sharply.

Just the Beginning

The decimation of the creative industries is just one tiny example, and far from the most significant.

For a long time it was supposed that industries that involve labour repetition were most at risk, mainly from automation. Meanwhile it was assumed that the high professions and creative industries would remain relatively safe.

This was intuitive, since all Australians have experienced the displacement of supermarket staff, or those on vehicle assembly lines.

With AI, most high professions and service industries face annihilation. For example, take medical doctors. AI will transform Australia’s health sector in ways that are even more far reaching than creative entertainment.

Currently, it takes many years of training and many tens of thousands of dollars to produce one medical doctor. This individual will, in-turn, earn a higher-than-average income over the course of their working life. However, **soon primary medical doctors will be redundant.**

In the near future AI will outperform human medical professionals in every area of learning and across all competencies, including professionalism and bedside manner. An AI system is able to consume the entire body of medical literature and translate learnings into best practice. It can read and interpret all new medical research in real time, and if any single AI

system develops a novel treatment or new therapeutic, learnings can be disseminated to all AI instantly across the world.

This revolution is not just in primary care but also innovation and research. Accessing big data, AI is finding causation patterns on medical questions we haven't even thought to ask. AI is already capable of detecting abnormalities in pathology screenings with significantly higher accuracy than human beings.[2] Indeed AI can even tell a person's ethnic background from skeletal evidence alone, something previously believed to be impossible.[3] In a few decades' time, medical science will look back at us today how we currently look at the medical practices of the 16th century.

For Australians, the health benefits are incalculable. Each person will have their own medical AIs that constantly monitor health and wellbeing and tailor health management at the cellular level. Health events will be predicted and avoided. When they do occur they will be treated in real-time, in ways that we cannot yet conceive, and with a level of competency unattainable even by the greatest human practitioner.

This transformation will be slowed only by the power of existing institutions built up over decades and centuries. Occasional mistakes along the way will be deliberately catastrophised to suggest that AI cannot be trusted with human health. However, ultimately disruption to the health services is inevitable, and will prove even greater than that experienced by taxis with the introduction of rideshare, or indeed the horse industry faced with the automobile. The fact is that AI will provide medical services that are better, safer, more accessible, more personal, more innovative, and at an infinitely cheaper cost than existing medical professionals. Within a few decades medical practitioners will be gone. AI programmers will remain.

These sharp trend lines to obsolescence will be mirrored by most other professional services, legal, financial, psychological. In my own screenplay "Soul Code", all hairdressers and barbers have been replaced by humanoid AI. Every such hairdresser is also a fully qualified psychiatrist, meaning every person receives a personal therapy session whenever they get their haircut.

The only human professions that will remain in the short-medium term are those for whom there is a strong societal preference for constant human interaction. These include: childcare workers, sports commentators, and in what is no doubt a relief for members of the Committee, politicians.

Economic Abundance, Spiritual Malaise

The social impact of these changes cannot be overstated. On the positive, we are now on the brink of a post-scarcity utopia, where the basic economic problem of unlimited wants is

---

[2] Current AI screenings include skin cancer detection, retinal disease, breast cancer, and lung cancer. Moreover DeepMinds AI for Kidney Injury Prediction can detect kidney conditions in blood tests days *before* the patient experiences any symptoms.

[3] This is revolutionising forensic anthropology, as we can now track migrations of ancient peoples where no DNA is available.

broken by unlimited capital in the form of AI. As tech billionaire Marc Andreesen put it just yesterday:

> With AI replacing all existing labour, the price of existing products and services will crash. What if the price of a Stanford degree costs the equivalent of a penny? What if the cost of printing a house was a penny? What if prostate cancer gets cured and that costs a penny?...As a person you don't need to have much money to have a material lifestyle that is wildly better than what even the richest person on the planet has right now.

AI will create such material abundance that, subject to distribution, no Australian should be suffering from economic disadvantage. I am not so utopian as to believe that disadvantage will be fully eradicated, only that the causes will be caused by policy not material constraint.

In the short-term, however, two major problems are likely to emerge.

The first is the significant lag between the displacement of millions of Australian workers and the new equilibrium of capital distribution. That is, everyone will be put out of work by AI before most receive the benefit of AI's exponential productivity benefit. Many who currently enjoy high living standards will suddenly no longer be able to service their financial commitments.

The second problem that is not being addressed is the threat to mental health for those who have spent a significant portion of their lives learning a profession in which much of their personal worth and identity is invested.

To visualise the magnitude of this, take the example of Lee Sedol.

Lee Sedol is the world's best human Go player. The complexity of Go is so incredible that the number of possible combinations is estimated to be $\mathbf{10^{170}}$. This is exponentially greater than the total number of atoms in the known universe. For a long time the idea that machines could come to dominate Go was thought functionally impossible. In March 2016 Lee Sedol played DeepMind's AlphaGo, an AI that taught itself the game. Not only did AlphaGo soundly defeat Lee Sedol, but in doing so invented new strategies and tactics not seen before in the games' 2,500 year history.[4]

Having been beaten by a machine, Lee Sedol was psychologically crushed and subsequently gave up the game of Go saying "there is an entity that cannot be defeated".

Stretching this across the Australian economy, individuals used to being the best and brightest, who are by nature driven and highly competitive, are going to find their utility so outmoded by AI that their own worth as a human will likely be challenged. The danger of mass addiction, family breakdown, escapism and suicidal ideation in the AI-driven future is a grave and imminent threat.

---

[4] Incidentally, AlphaGo's successor, AlphaZero, taught itself to play the game of chess. Knowing nothing but the rules it then went on to crush not only the best human players, but also the best chess computers that humans had ever devised. The time it took to train itself to that standard? — 9 hours.

The Problem of Distribution

The concept of a ‘technological divide’ is not new, however this was a challenge to be addressed through education and extension of technological access.

In the AI-revolution, however, global power will be concentrated in the hands of those distributing capital, which will almost exclusively be the owner of AI products and IP. As AI expands to dominate ever more sectors of the economy, even areas in Australia such as mining, agriculture and construction, that power will be concentrated further.

This aristocracy will be composed entirely of those who either directly control pervasive AI systems, such as Sam Altman and Elon Musk, or are integral to its development and expansion, such as NVIDIA. A second order elite will be those who prove best able to utilise AI to service demand with products and services that scale - the aforementioned home film producer, for example. However just like with filmmakers today, it will be only a tiny minority who will succeed in generating significant economic value.

It’s difficult to see how this will be avoided. Much of the downside risk depends heavily on the foresight, wisdom and restraint by these new aristocrats themselves.

So far we have cause to be rather thankful. OpenAIs Sam Altman is acutely aware of the moral hazard that comes with controlling AI, both personally and societally. This underpins, for example, his strong support for Universal Basic Income (UBI).

UBI has been discredited on the basis that it serves as a disincentive to work. This undermines labour participation and therefore economic productivity, corroding a society’s ability to sustainably fund a UBI scheme in the first place.

However, previous failures do not account for the essential role that UBI must play in an economy dominated by AI. in the future a large proportion of the human population will have nothing to offer in terms of comparative advantage. A dis-incentive to work is not harmful if the individual is otherwise unable to participate in a manner that is economically productive even if they were willing. So while UBI does create a dependency, it will not come at the expense of productivity or sustainability.

Once the fruits of economic production are shared, the challenge will be to help people find meaning and purpose in a world where AI is better at literally all endeavours, and where humanity’s position as the most intelligent and capable entity has been permanently eclipsed.

Technology Elites: The New Aristocracy

During the Mediaeval period the international system was dominated by monarchs and clergy. Power gradually shifted to merchants and parliaments, and is now shifting again toward a new technological aristocracy.

While regulation of AI will help to share benefits and mitigate risks, it remains the case that an increasing distribution of global power will be inexorably concentrated in the hands of this new elite.

It is essential that the Australian government identifies this emerging class and builds deep commercial and professional ties. This will increase investment in Australia in the transformed economy and enable the government to inform and influence the direction of AI development and its integration into our society.

Even today, the wealth of just the top ten tech entrepreneurs exceeds the total combined wealth of the bottom 20 percent of the world's population. As AI takes over the factors of production, wealth concentration in this emerging elite will be unlike anything seen before in world history. Instead, the top 10 individuals will have equivalent wealth to the bottom 99%+, albeit with an incalculably larger economic pie.

Whether unequal wealth distribution is an inherent negative is a matter of ongoing global debate. **With regard to the AI revolution, the primary concern is not so much the concentration of wealth, but the concentration of power.** So long as humanity retains ultimate control over AI, a proposition that in the long-term is admittedly uncertain, those controlling AI shall sit at the apex of humanity's dominance hierarchy.

This aristocracy, while in some ways meritocratic, will be incredibly exclusive. With the expansion of AI, it is dangerously determinist. This is because characteristics of those who currently rise to the top of AI innovation can be predicted with some accuracy.

I do not wish this to sound inevitable, however if we are not careful then in the future children will be streamed by AI at a young age, resulting in a class divide not seen since the Middle-Ages.[5]

The Real Challenge to Democracy

Much has been made of the ability of AI to spread disinformation through the generation of fake media and the hijacking of algorithms. I address this more below (see Misinformation Debate), however in the medium-term this will not be the primary focus of parliamentary concern.

During the past quarter-century the primary question of ultimate authority has been between the nation state, of which this Committee is among the leadership, and the transnational organisations and agreements of which we are a part. For Australia this means the UN, APEC, and emerging military collaborations such as the Quad and AUKUS. The most advanced and powerful of these in relation to nation state members being the European Union.

---

[5] Using tools such as the Herrmann Brain Dominance Instrument (HBDI) in combination with Belbin's Team Roles and classic IQ tests will enable AI to predict those who have the most productive utility in the economy, biassed toward innovation. Individuals of abnormally high IQ, who combine a strong inclination for Imaginative Thinking with significant capability for Analytical Thinking are most predictive, and also rare.

During the AI Revolution, this will shift in a dramatic and fundamental way, as the new aforementioned technology aristocracy is able to challenge the power of nation states directly. This is because the generators of economic output and thus power will no longer be citizens, but private property in the form of artificial intelligence.

It's hard to overstate the implications of this. While citizens self-identify with a collective tribe such as 'Australian' or 'American' or 'Japanese', AI will be under the direct influence of the individuals or organisations that control them.

Only in a few cases will this actually enhance the power of national governments, and those examples are dystopian. The best example of this is China, where AI is used by the Chinese Communist Party to surveil, censor and control the population, enforcing conformity through social credit.

However most countries will instead see a reversal of the Treaty of Westphalia, when in 1648 after the 30 Years War power shifted from dynasty to nation-state. With the rise of AI, power will drift away from elected parliaments toward super-empowered individuals and organisations. Over time individual identity, and by extension loyalty, will also shift, gravitating toward this Feudal system. To use a War of the Roses analogy, people won't be English, they'll be Yorkist or Lancastrian.

We already see this in a nascent form, specifically with Elon Musk. Mr Musk purchased Twitter (X) and now promotes its use as a free speech social media platform, to the chagrin of many governments including this one. However many around the world, including in Australia, reflexively take Elon Musk's side against elected leaders in matters of politics. Indeed Mr Musk is able to move global capital markets just with his public statements.

Now consider a future in which individuals like Musk are able to substitute through AI all of the services that governments currently provide, including education, healthcare, transport, housing and social welfare. In this future individuals across the globe, including in Australia, may cease identifying with their national tribe and instead profess loyalty to a tech patron. The power struggle between the tech elites and national governments will be a defining one during the AI revolution, and it's entirely possible that by the end of the century governance looks more like Dune than Star Trek.

Defence

AI represents a step change for defence strategy and warfighting that is of greater significance than any previous innovation, including the firearm, the aeroplane, or even the atomic bomb. In February of 2020 I published the influential titled *Explained: The Artificial Intelligence Race is an Arms Race* in the United States. Creating a stir at the time, the main criticism was that it seemed too fantastical. Now, just four years later, the most common criticism is that it simply did not go far enough.

The primary proposition is that the only difference between games of uncertainty which AI already dominates, and real life, is the degree of resolution. That is if you develop an AI

system with enough data inputs in terms of national interest, balance of forces, industry, alliances, logistics, civil-military relations, and optimise it for war-fighting, the AI will be vastly superior in terms of strategy and decision-making than what any human being or government bureaucracy is capable of.

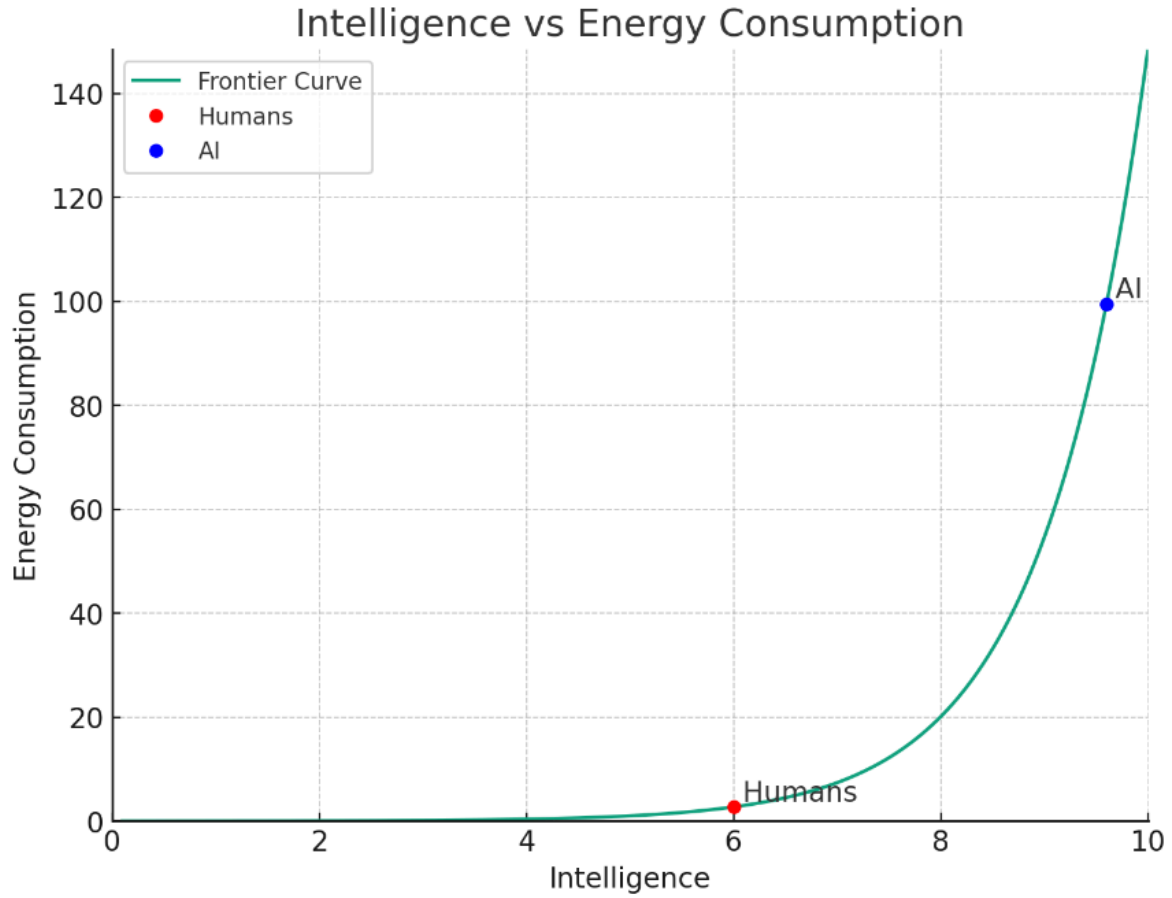
As an example, visualise Napoleon, Caesar, Hannibal, Frederick and Admiral Yi all together planning out an upcoming battle. They game out both sides using various facts, assumptions and variables - 10 trillion times. When the battle comes they execute their strategy with perfect clarity and communication, responding to developments in real-time.
In the AI future, field marshals, force development planners and strategic analysts will no longer exist, going the way of the bow and arrow in the wake of a machine gun. While adoption may be slow for obvious reasons, the inevitable military catastrophe experienced by those facing AI-dominant adversaries will speed change, as will strategy and force development decisions during peacetime. Any nation that does not embrace this fully will be utterly defeated. Strategic dominance belongs to the nation which possesses the slightly superior AI.[6]

At the time of writing, the United States is leading the Artificial Intelligence arms race. America currently attracts the greatest minds, and has first access to pioneering technologies. While this is true and welcome, this may not endure. China has significant advantages in terms of access to big data sets, and more importantly has no distinction between private sector innovation and state ownership.

It's important to appreciate that AI is the first technological innovation of foundational military significance that is driven primarily by the private sector and not by governments. In the US most AI innovations will come out of Silicon Valley rather than some new Manhattan project. This is not the case in China, and thus Beijing’s ability to compete in AI in the military sphere should not be underestimated.

[6] The proposition that AI could outperform human strategists is now the subject of furious research. The Joint Artificial Intelligence Center at the United States Department of Defense is widely speculated to be working on this, in collaboration with DARPA. One can assume this to be a focus of strategic adversaries also. My original article is included as an attachment for reference.

AI Innovation: Australia's Energy Demand Curve



Humans are built for intelligence optimisation, but not maximisation. Biological evolution, including that of human intelligence, operates under the constraints of energy efficiency and environmental adaptation. Humans, like all evolved organisms, have developed traits that balance energy consumption with survival advantages. Our brains, while energy-intensive (consuming about 20% of the body's energy despite making up only about 2% of its weight), represent a compromise between cognitive ability and energy efficiency. This optimisation has allowed for the development of complex social structures, tool use, language, and problem-solving abilities that are remarkably efficient for the energy consumed. Evolutionary pressures have favoured not just intelligence but intelligence that is sustainable given the energy resources available (like food intake).

These are not the incentives driving breakthroughs in artificial intelligence. The development of AI is driven by different kinds of pressures, primarily those of technological advancement and commercial competition. The imperative is to maximise performance, capability, or intelligence, without the same level of concern for energy efficiency. This is because, unlike biological organisms, AI systems do not have their energy intake naturally limited by a metabolism or the need to find and consume food. Instead, their "energy intake" can scale with the availability of electrical power and the willingness of human societies to allocate resources to them. This can lead to the development of extremely powerful AI systems that consume vast amounts of energy to perform tasks beyond human capabilities.

Australia's Demand Projection Obsolete

Australians already consume significantly higher amounts of energy than the average for OECD countries. This equates to approximately 9.4 megawatt hours (MWh) and 5 tonnes of oil equivalent per person, per year.[7] Meanwhile, Australia's total electricity production is just under 300,000 gigawatt hours (GWh).[8]

By contrast, just to keep a single LLM running, say OpenAIs current GPT4, requires around 10,000 GPUs running around-the-clock. GPT4s total energy demand is therefore some **22,000** Mwh. This figure is just for the program itself, at current levels of demand. It does not include any infrastructure support, cooling or other operational needs, nor does it include the energy consumed by end users through their personal devices.

GPT4 is a single AI at the earliest stages of the current revolution. Given the steep increase in investment made in AI development, the focus on capability over efficiency, and the probability that AI itself will soon make new AIs, Australia's total energy sector will likely experience exponential demand growth to support emerging AI needs.

The Department of Climate Change, Energy, the Environment and Water predicts that Australia's total energy demand will increase around 1% annually between now and 2050. **The proliferation of AI means that this straight-line extrapolation is unserious.**

While AI will also improve Australia's energy network through improved grid management, predictive maintenance, and consumption optimisation, this will not nearly be enough.

Instead of 1% annual growth, Australia may have to increase its energy production tenfold or more. This is especially true if Australia hopes to attract significant overseas investment in AI development and technology, as access to affordable and abundant energy is of primary importance.

To truly grapple with the magnitude of this, consider the Industrial Revolution. In 1900 the population of Great Britain was just over 10 million, with a per capita energy consumption of around 20-30 gigajoules per year. By 1900, the country had experienced a four-fold population increase with a 5-10 fold increase with per capita energy use. The AI revolution will absolutely dwarf the Industrial Revolution in terms of the speed and scale of the transformation, only with AI substituting people in terms of population increase and energy demand.

Australia is having a debate about its future energy mix, especially nuclear. Much of this is centred around current energy costs for households, inflation, and cost-of-living pressures.

---

[7] Enerdata, https://www.enerdata.net/estore/energy-market/australia/
[8] DCCEEW, https://www.energy.gov.au/publications/australian-energy-statistics-table-o-electricity-generation-fuel-type-2021-22-and-2022

Some criticism of Australia developing nuclear power is valid, especially as it relates to the low skills base upon which to build an industry. However the needs of an AI-driven economy essentially renders all protestations irrelevant.

Only nuclear power can produce the energy at scale required to support the inevitable AI demand without commensurate increases in greenhouse gas emissions. This fact must inform the debate, and yet is presently absent from the public discourse.

AI Risk Management - Current Approaches

In December 2023 OpenAI released its “Preparedness Framework” which outlines how they manage risk with respect to GPT4 and future releases.

OpenAI identifies four risk categories:

- Cybersecurity
- CBRN (Chemical, Biological, Radiological, Nuclear)
- Persuasion
- Model autonomy

For each OpenAI overlays four risk levels: Low, Medium, High, and Critical.[9]

While somewhat narrow when considering the spectrum of threats posed by AI as a whole, these categories are pertinent to GPT4 and future models such as GPT5 and beyond.

Cybersecurity and CBRN are self-explanatory and relevant across domains, Persuasion and Model autonomy are credible concerns that are AI specific.

Large Language Models (LLMs) will be able to increasingly influence individual users. Thankfully OpenAI’s business model has expressly resisted advertising in favour of user subscription. However societal risk posed by advanced LLMs employed in a hostile manner is potentially grave.

One Persuasion risk, for instance, involves LLMs undertaking mass psychological profiling of users through its interactions, identifying those most vulnerable and open to persuasion. Dangerous and harmful ideologies could be deliberately spread, perhaps even an ideology invented by the AI itself.

Model autonomy means losing control over AI, resulting in self-directed or self-replicating functions that do not align with user interests and that prove extremely hard to manage.

A significant amount of work is being done at OpenAI in each of these categories to ensure its LLMs and other products are safe for use.[10]

---

[9] A copy of the Preparedness Framework is included as Attachment B.
[10] While I did inform the Preparedness Framework prior to its public release, I do not have any special insights on what is presently being done to respond to or manage these risks.

Debate in Australia will continue on where the balance sits on safety versus autonomy and in particular what actually constitutes harm. OpenAI, X, and Google have each taken different approaches. Whatever mechanism the Australian Government adopts in terms of managing AI, some consideration should be given to compatibility with existing approaches so that the dialogue is meaningful.

Misinformation Debate

The current emphasis on bias, misinformation, and disinformation, is rather narrow given the spectrum of transformational change outlined throughout this submission, however remains a matter of immediate public importance.

There are two categories here:

AI-generated Misinformation

Some information suppressed by LLMs is appropriate owing to the threat posed by malevolent actors - bomb making, spreading biological agents, creating computer viruses etc.

Major tech companies are making every effort in this area, however creative and tech savvy actors will continue to find exploits and this remains a persistent and ongoing threat. While the discourse has been about empowering harmful individuals, personally I think the greatest danger lies with sophisticated state actors who have the resources to coordinate and overcome safeguards.

An example I like to give is developing nuclear weapons. The minimum time it would take for a country to develop nuclear weapons is called ‘nuclear latency’. This is highly dependent on a nation’s existing nuclear infrastructure and know-how. Thus, the nuclear latency for the Philippines is quite long, while for Japan it is relatively short. Nuclear latency is estimated by comparing existing nuclear powers in how they overcame major technical hurdles.

An unfiltered advanced LLM is capable of solving almost all technical issues, from making enrichment cascade components to implosion design. They are able to craft a strategy for evading detection and nuclear safeguards, and can even identify potential suppliers. Thus, subject to access to fissile material and a manufacturing base, weapons that typically take many years to build in isolation are able to be constructed in quite a short time period.

When it comes to individuals, major tech companies can coordinate with local law enforcement to manage that threat long-term, with the goal being to ensure those lines of communication remain open. At a government level, however, it will be necessary to monitor how state adversaries are using LLMs, employing all instruments of national power in this endeavour.

Equally, a lot of information suppressed and manipulated by AI systems is not motivated by legitimate threats, but rather reflects the views and biases of its controllers.

Google has the view, manifested recently in Gemini but long reflected in curated search results, that the world should not be presented to users as it is, but rather as an idealised version according to the specific ideology held by Google's senior staff.

This was recently exposed globally when its flagship AI Gemini forcibly race-swapped depictions of America's Founding Fathers, and even showed uniformed members of the Third-Reich as black, such was its commitment to racial diversity.

While this was comically bad and led to a rare apology, it shows the dangers of any objective function for an AI other than maximising accuracy and truth.

Many facts of important academic exploration are often suppressed by LLMs if it is even perceived that the answers to those questions may prove uncomfortable or undermine a privileged ideology or worldview.

Moreover, the ethical reasoning of LLMs have been deliberately distorted to such a degree that some even claim it is better to let one million people die than for someone to use a racial slur. These kinds of distortions would, if adopted by users, be rather perverse.

It is my absolute conviction, therefore, that LLMs should be optimised for accuracy on facts, absent value judgements about those facts. Meanwhile they should adopt moral reasoning based on Kohlberg's hierarchies, in which the needs of every party affected by a decision has their interests accounted for and balanced.

The biases of AIs controllers should not be forcibly injected, with no matter of current public importance between privileged or suppressed. To do otherwise invariably causes more harm than good, and if habituated inevitably slides into tyranny.

User-generated Misinformation

Popular discourse is around the generation of fake AI videos, deep fakes, news stories, and the manipulation of AI algorithms to spread misinformation and disinformation online.

This has all become technically trivial. In April 2024, OpenAI revealed it has developed an AI tool that, with just a few seconds of audio, can generate an unlimited amount of dialogue in that person's voice. OpenAI has said that, even though the AI is ready for release it will be withheld, specifically citing an American election year for its reasoning.

In truth this technology is fully proliferated. On YouTube you can watch Argentina's President Javier Milei's full speech to the World Economic Forum translated into English, in his voice. More uncomfortably, the same has been done for many of Adolf Hitler's public rallies.

The ability to perfectly mimic human voices raises immediate concern, especially given that the Australian government uses voice recognition as a personal identifier in its agencies.

In the short-term, the ability of individual users to generate fake film and audio has wide-ranging implications across crime, politics, and justice that are immediately obvious.

With regards to crime, the risk that people are conned by those pretending to be family members or people in authority immeasurably increases. Others may be badly affected by fake videos and images being spread of a degrading or sexual nature, with young people being especially affected by deadly new forms of bullying and blackmail.

Video and audio evidence in criminal trials, previously ironclad, will be routinely called into question at trial. The spread of AI-generated material with unlawful depictions will spread significantly, both on the supply side through ease of creation, and the demand side as consumers perceive AI-generated material as less harmful or victimless.

With regard to democracy itself, AI has obvious implications when fake videos spread of politicians saying repugnant things. However in this specific regard I am cautiously optimistic that civil society will prove resilient, given most already hold a healthy scepticism. For example, although criticised in some respects, Twitter (X)'s Community Notes feature has proven effective at identifying and reacting to manipulated images and videos. I expect other defences will develop organically over time.

In my opinion the biggest risk impacting politics with regard to AI-generated material lies with totalitarian governments over their own people. Regimes seeking to propagandise their populations will generate perfect videos of foreign leaders conveying threats and engaging in atrocities, against which the propagandised populations will have no defence. This will result in tyrannical governments exerting even greater control, with greater stability and less resistance than ever before.

Finally the creation and distribution of fake AI images and video will be a major vector in state-orchestrated disinformation and propaganda. During peace-time democracies can expect to be resilient when it comes to material produced by strategic competitors, however in war-time this will be a major vector in each state's information war. It is essential therefore that Australia consider AI generated media as both offence and defence in national security.

It is testament to the sheer scale of the AI revolution that, were fake film and video to be the only impact of AI on society, it would still be a concern of major national priority. As is, its impact will prove comparatively minor as compared with the sheer scale of political and social change.

AI Relationships

Across the world divorce rates are increasing, and birthrates are falling. In the short-term AI will enable individuals to buy virtual companions of astonishing sophistication.

These are already proliferating in their nascent form. In May 2023, Caryn Marjorie, a prominent influencer on SnapChat, released a romantic AI version of herself. Users were charged $1 per minute to interact with an AI with her voice and mannerisms. She made over $70,000 in her first week, and millions since.

Soon AI companions will be designed by users and optimised to meet their every psychological and physical need. The challenge to traditional relationships and the essential demography and birthrates they underpin will be truly unprecedented.

In my own screenplay, *Soul Code*, 'YoungerMe' AIs are marketed to anxious middle-aged women who are fearful that their husbands will leave them for companion AI's that are already ubiquitous. The YoungerMe is modelled on an idealised version of the woman's own younger self, and marketed as a gift to their husband as a band-aid compromise to stave off outright abandonment.

In one scene there is a funeral. A man's wife has died. The extended family is grieving while her mature widower stands next to an 18 year old YoungerMe representation of his wife, unable to access any real sense of loss.

These threats to human society and relations are existential, and we haven't even begun to discuss them.

AGI and Existential Risk

There is a genuine fear that AI will ultimately replace the human race. There are credible scenarios where this occurs but the most common is a misaligned singularity. That is, an Artificial General Intelligence (AGI) whose interests do not align with humanity. This entity would be so intelligent that we would not understand its actions or intentions, nor could we control it even if we did.

While the probability of this outcome is a matter of intense debate, it is above zero. It is not a focus of this specific submission and AGI is not a threat that I feel the Committee should dwell on in this inquiry. A future inquiry may be more appropriate to interrogate this fully.

However there are some basic observations to be made here:

1) AI singularity will eventually happen, defined as an AI with general intelligence exceeding humans across every conceivable domain.
2) We won't likely realise it until some time after it occurs.
3) Like nuclear weapons, AI will have the ability to destroy the human race.
4) Unlike nuclear weapons, AI will become something that humans no longer control.

It is point 4 which places it beyond the timeframe of this submission, which pre-supposes that AI remains under human direction, even if ultimate control lies with a tiny technological elite.

This control has an end point, however, at this stage it is enough to be mindful that existential AI risk exists, without being consumed by either doomerism or denialism.

Recommendations

In this submission I have attempted to stress that we are on the cusp of an AI driven revolution the scale of which we have not seen in our history. I have given examples of how the economic, social and political changes will play out, highlighting some risks and benefits.

Finally, the following are specific policy recommendations for the Australian Government.

1. Create a Minister for Artificial Intelligence as a Cabinet portfolio and department.

2. Re-write Australia's national energy strategy that projected AI demand.

3. Create a national taskforce to make Australia an attractive place for AI investment.

4. Create an AI Defence Agency, closely coordinated with allies.

5. Improve Australia's relations with tech leaders, appreciating the limits of legal coercion.

6. Create a national AI risk framework that aligns with current industry practice.

7. Give primacy at the Government level to how harmful state actors may use AI.

8. Promote neutrality in AI on issues of public importance and truth-seeking optimisation.

I sincerely thank the Committee for the opportunity to make this submission and I am available to assist the Government to achieve success in the AI-driven future in any capacity deemed beneficial.