



Supporting Responsible AI

Future For Now Submission

July 2023

Australia's AI Ethics Principles provide a solid foundation, but there is room for improvement. Currently, the principles mainly concentrate on the AI systems themselves, with little depth on the application of AI by businesses. While the AI system developers may adhere to these principles, it is not clear how this extends to organisations using AI.

Take, for instance, the principle of "Contestability." Although it may be upheld in some contexts, the scenario of organisations replacing workers with AI. While this move may boost a company's profitability, it raises questions about the rights of the affected workers and their ability to contest such decisions.

Given the widespread accessibility of generative AI to both organisations and non-technical individuals, it is imperative to revisit and revise the AI Ethics Principles.

The focus of this submission by Future For Now, is on the application of AI by businesses, rather than the building of LLMs and ADM tools.

Bridging the Digital Divide Through Education

We are not strangers to industrial revolutions; the unique challenge we face now is the unprecedented speed at which this revolution is unfolding due to the world's extensive, existing digital infrastructure. In contrast, previous revolutions powered by technologies like steam engines, electricity, and the internet required the infrastructure foundations to be built over extended periods, allowing ample time for the workforce to transition.

While the conversation often highlights AI as a technology that enhances human capabilities rather than replacing them, the harsh reality in a corporate environment is that there will inevitably be both winners and losers.

The government has a crucial role to play in developing a comprehensive roadmap, ensuring AI's accessibility to all and preventing an unintended consequence of an exacerbated digital divide. This roadmap must encompass strategies for re-skilling individuals and re-tooling businesses.

Consider the potential risks associated with AI - concerns over robustness, reliability, bias, and even hallucinations. Without an in-depth understanding of these issues, a workforce devoid of AI literacy will struggle to adapt their roles effectively to mitigate these risks, ultimately resulting in the uneducated falling behind.

The deployment of AI cannot be viewed in isolation as a mere technological event; its implications are unmistakably social and economic. To genuinely guarantee that all Australians benefit from AI, the Australian Government must prioritise subsidised training for the workforce as a fundamental component of responsible AI application.

Government-subsidised training for the Australian workforce is a cornerstone of responsible AI implementation.

We must expand the scope of STEM education to empower a workforce proficient in analytical, creative, and adaptive communication skills, alongside critical thinking. These competencies will enable workers to safely and effectively enhance their job outputs.

Here are some areas for consideration:

1. **Curriculum Reform for the Future:** Reforming the state-based school curriculums to reflect that the future of coding may be the English language itself. With AI's advancing capabilities in natural language processing, students will need less learning on traditional coding languages and more focus on foundational communication skills for an AI-driven future.
2. **Adult Learning and Transition Support:** Adult learning programs are key, not just retraining, but also transition support for those whose roles could be displaced by AI and automation. Roles based on repetitive tasks and simple cognitive functions are already being replaced by autonomous agents. The government might need to support a shift from traditional professional roles to those supporting community needs, an aging population, and other under-funded and under-represented areas where human skills are desperately needed (eg. teaching and nursing).
3. **Empowering Small Businesses in an AI Era:** Small businesses can benefit from regional workshops that equip them to compete in an AI-driven marketplace. Government support should ensure equitable access to AI technologies through tax incentives and provide training on their responsible application including - strategic planning on AI application, risk management, growth opportunities, and workforce reskilling.

In short, training and education for individuals and small businesses must emphasise more than just technical applications. It should focus on fostering critical thinking, AI literacy, and understanding.

Ensuring Diversity & Reducing Bias

Generative AI technology, powered by LLMs, will play a significant role in our digital landscape. However, as mentioned in the *Safe and Responsible AI in Australia Discussion Paper*, it is crucial to recognise that these LLMs are susceptible to biases stemming from both the underlying datasets, such as underrepresentation, and the additional training through Reinforcement Learning from Human Feedback (RLHF). Biased representations in the training data can perpetuate social stereotypes and result in unfair discrimination when the language model makes predictions. This poses a significant challenge, as AI becomes increasingly autonomous and integrated into our daily lives.

Addressing bias in AI is paramount to promoting diversity of opinion and experience, preventing the fragmentation of society. If we fail to tackle these biases, we risk perpetuating harmful echo-chambers, as seen through social media algorithms. Learning from past mistakes, it is imperative to limit biased models and hold big-tech accountable for the societal impact of their AI technologies. We cannot afford to let the future be controlled solely by a select group of individuals.

By adopting an inclusive and thoughtful approach, we can work towards creating AI systems that truly benefit everyone, fostering a more equitable and harmonious society. To ensure this, some strategies could include:

1. **Establishment of an AI Diversity Forum:** The Australian Government should establish a diverse forum comprising people from various backgrounds to actively shape policies, processes, and technology adoption related to AI. By incorporating a wide range of perspectives, including gender, race, neurodiversity, ethnicity, age, sexuality, worldview, and skillset, we can foster a more inclusive and ethical approach to AI implementation. Embracing the wisdom of the crowd will be instrumental as we navigate this new age of technology. Such a forum could potentially be hosted via a Decentralised Autonomous Organisation (DAO) and handled through blockchain technologies. This would allow members of the forum to vote on policies, processes and technologies around AI adoption.
2. **Adopting or establishing bespoke benchmarks to assess LLMs:** The Australian Government could consider adopting or creating specific benchmarks, such as the 'Bias Benchmark for QA' (BBQ) and the 'Bias in Open-Ended Language Generation Dataset' (BOLD), to assess the bias in LLMs. These benchmarks would track the performance of LLMs, ensuring that they do not perpetuate biases. By using these established datasets as a foundation, Australia can ensure the development and deployment of future LLM models that prioritise fairness and inclusivity.
3. **Building a Proprietary Australian LLM:** In order to avoid overreliance on internationally-created LLMs and preserve national identity, the Australian Government should explore options to fund and develop a proprietary Australian LLM. This LLM would be tailored to the unique characteristics of Australia, encompassing its ancient and modern history, Indigenous culture, laws and regulations, colloquial language, and diverse cultural nuances. Training such an LLM with a focus on ethical considerations, such as compliance with the Modern Slavery Act, intellectual property, copyright laws, and embracing diversity, will be essential to ensure responsible AI implementation. Such a LLM could then become the backbone of the public sector, be made accessible to the Australian public, as well as other AI systems.

Fostering Trust Through Transparency and Disclosure

The responsible use of AI necessitates transparency across the entire value chain, from the language models themselves to their integration layer, the businesses utilising them, and ultimately, the end consumers. Although AI systems may adhere to Australia's AI Ethics Principles, accountability tends to dilute as we move through business applications to the end consumer. We propose government attention in three key areas: Language model disclosure, interaction transparency, and monitoring guidelines.

1. **Language Model Disclosure:** Numerous interface providers, app developers, and SaaS companies are integrating AI into their products. It's crucial to have a traceable accountability path built into this application layer. These companies should be required to disclose the existence and characteristics of any language models functioning behind their user interfaces, helping businesses make better-informed decisions about their tech stack and AI decision-making processes. This could be extended to being a consumer facing disclosure, modelled on existing Privacy Disclosure Statements.

Moreover, we need guidelines to assist small businesses in making informed choices about AI system implementation, addressing questions such as ownership of the language model, its creation process, training datasets, and model security against potential hacks or changes.

2. **Interaction Transparency:** Consumers have the right to know when they're interacting with an AI system. Similar to voice recording laws, users should be informed when dealing with AI. This awareness empowers consumers and businesses to make informed decisions about AI usage, fosters healthy scepticism, holds AI systems accountable, and can help with the acceptance of unexpected responses (like AI hallucinations).

For instance, 'AI in the room' disclosures should be made when AI is involved in tasks such as note-taking or transcribing conversations during a meeting. Similarly, if a customer service representative is an AI powered chatbot, not a human, it should be clearly communicated.

3. **Monitoring Guidelines:** Rapid integration of AI into business workflows presents challenges as AI often functions as a 'black box.' It can be difficult for a business to monitor potential issues like toxicity and hallucinations, especially at scale.

Regulations could mandate AI technology to enable end-user feedback reporting in a measurable manner. This would be a crucial tool for consumers and businesses alike to build trust and confidence in AI's capabilities. It would also close the loop with the bias benchmarking recommended above.

Trust stems from explainable, accountable value chains and respect for human intervention within business workflows.

Taking a visionary approach to the future

Practically, the aim is to establish guidelines and share best practices that allow businesses to balance safety while embracing the opportunity for innovation.

It is vital that the Government not only address tactical aspects of AI application, but also adopt a more far-sighted approach. The Government's current Responsible AI Discussion Paper includes some practical but narrow perspectives on the impacts of AI. It is crucial to step back and examine the larger picture, considering the rapid pace of revolutionary change, its impact on the workforce, and the ensuing social and economic disruptions.

A visionary approach could envisage what an AI-empowered life would look like for Australians a decade from now. It would entail contemplating new economic frameworks, such as how to tax businesses that operate with leaner workforces, and how to reinvest resources in pressing areas like climate change, aging population, and growing underemployment.

In the AI-driven future, we should strive not only for technological advancements but also for a society that uses AI to enhance human life and address the pressing challenges of our time.

About Future For Now

Future for Now is a collaborative project focused on helping businesses apply technologies such as AI, automation and low code solutions into their workflows.

We are not futurists. Whilst we do get excited by shiny things, we go beyond the hype and focus on the real world applications of technologies. This is anchored in recognising where business growth and efficiencies lie, whilst being realistic about the cost of inaction.

Our approach is to always anchor against business growth and driving efficiencies, whilst being realistic about the cost of inaction and the speed of change in an industry sector.

We recognise the impact that AI will continue to have on the workforce and are passionate about creating pathways toward careers of the future.

We do this in two ways, firstly by **guiding leaders** on digital transformation and organisational change. Secondly, by **empowering individuals** to work with tech that will enhance their job.

Get in touch



Lee Schofield

Co-Founder
Future For Now

E: lee@futurefornow.world



Simone Blakers

Co-Founder
Future For Now

E: simone.blakers@liquidcx.com.au



Find us on LinkedIn:

<https://www.linkedin.com/company/future-for-now/>