INQUIRY INTO ARTIFICIAL INTELLIGENCE (AI) IN NEW SOUTH WALES

Name: Dr Darcy W.E. Allen, Professor Chris Berg, Dr Aaron M. Lane

Date Received: 12 December 2023

Parliament of New South Wales
Legislative Council
Inquiry into Artificial Intelligence in New South Wales

Authors:

Dr Darcy W.E. Allen, Professor Chris Berg, Dr Aaron M. Lane

Contact Author:

Dr Aaron M. Lane, Senior Research Fellow, RMIT Blockchain Innovation Hub, RMIT University.

Date:

12 December 2023

Dear Committee Members,

We welcome the opportunity to contribute to this important inquiry into artificial intelligence (AI) in New South Wales. We are a team of academics from RMIT University. We are deeply involved in research on the digital economy through the RMIT Blockchain Innovation Hub and RMIT Digital3.[1] Our industry-engaged research examines frontier technologies — notably generative AI and blockchains — from the perspective of economics, law, and public policy. We have collectively spent decades studying, using, and building the technologies and business models of the digital economy. RMIT Digital3 also have a range of research-backed short courses including 'Generative AI: Implications and Opportunities for Business'. The breadth of our research informs our views on how generative AI provides opportunities for New South Wales.

Introduction and Outline of Submission

This submission focuses on Generative AI, a new example of a general purpose technology that can be applied and built on across a wide variety of NSW industries. While there is widespread understanding of the economic importance of general purpose technologies, identifying their precise applications and economic benefits is hard.[2] Partly this is because we simply don't know definitively what these technologies are for, and because the benefits are

---

[1] RMIT Blockchain Innovation Hub, https://rmitblockchain.io/
RMIT Digital 3, https://www.rmit.edu.au/partner/hubs/digital3
[2] On general purpose technologies see Bresnahan, T. and Trajtenberg, M. (1995) 'General Purpose Technologies 'Engines of Growth'?' *Journal of Econometrics 65(1)*: 83-108

diffuse across an economy and across time. While there is growing evidence that generative AI provides some productivity benefits, there are also a broad range of other potential benefits, such as the reduction of various transaction costs and the reorganisation of economic activity.[3] There is significant potential for generative AI to be applied in key industries in the NSW economy. For instance, we and others have laid those benefits out elsewhere, including in specific sectors, such as our submission to the House of Representatives inquiry into generative AI in education.[4]

While we are optimistic that generative AI will provide significant benefits to NSW, our approach in this submission is to focus on some of the main concerns that have been raised. There have been widespread calls to slow down the acceleration and application of generative AI.[5] The initial response to a new disruptive technology, let alone one that has human-like characteristics, is to apply a precautionary principle.[6] The precautionary principle around new technologies involves overweighting the current tangible harms of the technology, and underweighting the potential future benefits.

Our contribution in this submission is to address four main areas of concern that are often raised regarding generative AI. We structure our submission around the following frequently asked topics:

1. Hallucinations

2. Bias

3. Impact on work

4. Competition and platforms

Our aim is to provide evidence on how to think about these genuine concerns that people have about generative AI. Understanding and adapting to these concerns is critical if we are to enable the necessarily entrepreneurial experimentation of generative AI into NSW business models.

---

[3] Potts, J., Allen, D. W. E., Berg, C., and Ilyushina, N. (2023). 'Large Language Models Reduce Agency Costs'. *Available at SSRN.*

[4] See Submission 18, RMIT Blockchain Innovation Hub Researchers https://www.aph.gov.au/DocumentStore.ashx?id=f88c769c-c0a0-42da-a2e9-d5ca9aba7f89&subId=745186

[5] See 'Pause Giant AI Experiments: An Open Letter', Future of Life Institute, 22 March 2023, Available at https://futureoflife.org/open-letter/pause-giant-ai-experiments/

[6] Thierer, A. (2016). *Permissionless Innovation: The Continuing Case for Comprehensive Technological Freedom.* Mercatus Center at George Mason University.

1. Generative AI hallucinations

A common misconception about generative AI revolves around its intended purpose and functionality. Unlike traditional search engines designed for delivering accurate, factual information, generative AI operates as a prediction engine.[7] This key distinction underscores its primary purpose: fostering creativity rather than ensuring accuracy. As non-deterministic systems, generative AI models excel in creativity.[8] This creative ability propels their applicability across many new domains as a general purpose technology. But while the non-determinism of generative AI models is the source of their benefits, it also contributes to what are often termed as 'hallucinations' in their outputs. These are instances where the AI generates content that — while potentially unique, creative and even plausible — may not be factual.

The community of generative AI users must evolve their understanding of the technology to incorporate the potential for hallucinations. They might also seek to use additional features of the technology (e.g. the capacity to browse the internet) and other approaches (e.g. manual search, judgement, and expertise) to mitigate some of the risks. Indeed, there is already increasing awareness and scepticism regarding some of the outputs from these models. This scepticism is positive and welcome and is part of a healthy use of consuming and applying generative AI outputs in practice. As users grow more accustomed to the idiosyncrasies of generative AI, they are learning to navigate its capabilities and limitations. Greater user discernment is crucial and fosters a more informed and critical approach to the consumption of generative AI.

2. Bias in generative AI models

Generative AI models are regularly benchmarked against some idealised, perfectly unbiased AI. Signs of bias are quickly identified so that the technology can be slowed, or each bias can somehow be corrected. While the bias of generative AI outputs is a major and legitimate concern for many users and observers of the technology, this benchmarking approach is unrealistic and unhelpful. Such a standard is unattainable for several reasons, including the complexities of both technology and humans. Addressing this bias requires an understanding of the limitations of generative AI models, as well as a comparison to alternative tools or technologies to solve particular problems.

---

[7] Agrawal, A., Gans, J., & Goldfarb, A. (2022). *Prediction Machines, Updated and Expanded: The Simple Economics of Artificial Intelligence.* Harvard Business Press.

[8] On the non-deterministic nature of generative AI models, see: Berg, C., Davidson, S., and Potts, J. (2023). 'Institutions to Constrain Chaotic Robots: Why Generative AI Needs Blockchain'. *Available at SSRN 4650157*.

The biases in generative AI models are, in part, a reflection of the biases inherent in humans. These models are trained on vast datasets, often from accessible internet. Unsurprisingly biases from the datasets become embedded in the models. This is capturing the prevailing tendencies, preferences, and prejudices of the data it has been trained on. Attempts to correct for these biases, such as reinforcement learning through human feedback, might mitigate some of these biases, but it may also integrate new ones (from the humans in the reinforcement process).

The critical question for users of generative AI is not whether bias exists, but rather what types of bias are acceptable in specific contexts (i.e., biased compared to what?). Users must consciously navigate the trade-offs between different sources of bias: the averaged biases of input data in large language models (LLMs), biases introduced through safety and human reinforcement interventions, biases inherent in individual or group perspectives, and so on. This role of humans in the loop of generative AI is essential, it dictates the level and type of bias users are willing to tolerate and correct for in their interactions with AI.

Awareness and acknowledgment of these biases are crucial. Users must not only recognise the potential biases in AI models but also actively respond to them. This response may involve adapting our norms and behaviours in how we interact with these technologies. As we become more cognizant of the generalised biases present in AI models, our engagement with them must evolve. This approach is similar to dealing with the challenges of hallucinations as described above. This evolution in interaction and norm-setting is a vital step in responsibly harnessing the capabilities of generative AI, ensuring that its benefits are maximised while its biases are acknowledged and managed. Importantly, the approach we have laid out here differs widely from a model-level approach of assuming we can create some unbiased generative AI model itself.

3. Generative AI and the impact on work

There are significant potential productivity benefits from generative AI in the NSW economy. While estimates are necessarily predictive the early results suggest staggering productivity gains and improvements. For instance, a recent Harvard Business School paper studying the productivity benefits of consultants with generative AI (compared to those who completed tasks without) showed across-the-board productivity gains.[9] Indeed, as the author Ethan Molluck describes, "Consultants using AI finished 12.2% more tasks on average, completed

---

[9] Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., ... & Lakhani, K. R. (2023). Navigating the jagged technological frontier: field experimental evidence of the effects of AI on knowledge worker productivity and quality. Harvard Business School Technology & Operations Mgt. Unit Working Paper, (24-013).

tasks 25.1% more quickly, and produced 40% higher quality results than those without. Those are some very big impacts."[10]

Just because generative AI increases worker productivity does not mean that robots will take our jobs en masse. Unlike technologies that purely automate, generative AI applications typically require a process between a prompting-human and the technology. Generative AI is applied as a process of co-production. Human expertise is needed to craft effective prompts, and to identify valuable problems and applications that generative AI models might help with. Furthermore, effective co-production with generative AI typically involves feedback loops and responses from human prompters, including an almost-entrepreneurial process of making judgements over outputs and adapting to them.[11] This suggests a significant role for humans remaining in the process. Many of the productivity improvements through generative AI will come through replacing tasks not jobs.[12] Co-production is not merely about automating processes but enhancing them through a deep understanding of the nuances involved in each task.

Generative AI is fundamentally a bottom-up technology, requiring significant human input not only in guiding these systems but also in making critical judgments about their outputs. These are not standalone solutions but tools that need to be applied to a diverse range of tasks, each requiring deep, contextual local knowledge. The human element in directing and interpreting the outputs of generative AI is indispensable, emphasising the technology as an augmentative tool rather than a replacement for human skills.

4. Competition and platforms in generative AI

As in all sectors, competition is a crucial driver of innovation and growth in the generative AI sector. Competition in the generative AI sector manifests at different levels.

● Competition between the underlying foundational large language models.

● Competition between the general interfaces of those models.

● Competition in the "wrappers" and applications built on top of the models.

For instance, there is differentiation in products such as ChatGPT and Bing, which, despite being based on the same underlying model, offer distinct user experiences and functionalities.

---

[10] Molluck, Ethan (2023) 'Centaurs and Cyborgs on the Jagged Frontier' *One Useful Thing*, 16 September, https://www.oneusefulthing.org/p/centaurs-and-cyborgs-on-the-jagged
[11] Potts, J. (2023). 'The Use of Knowledge in a Digital Economy'. *Available at SSRN 4440771.*
[12] On the impact on tasks in an Australian context see Walkowiak, E. and MacDonald, T. (2023) 'Generative AI and the Workforce: What Are the Risks?' *Available at SSRN 4568684.*

This diversity in applications highlights the dynamic and competitive nature of the generative AI field.

The current regulatory debates around generative AI should consider the importance of these competitive dynamics. In the first instance, the current nascent stage of AI technology means that we may not fully understand the implications or the full scope of what we are regulating. This makes it difficult to regulate and leads to prominent calls for major changes such as licensing AI models.[13] Such approaches carry a significant risk of regulatory lock-in and capture. It is crucial for lawmakers and regulators to recognize that any regulatory impositions, such as licensing requirements, can inadvertently lead to a concentration of power in the industry, thus stifling competition.

We expect that the market for generative AI models and products to be diverse. It will consist of both large closed corporate models and a rich, diverse array of open-source models. This diversity is desirable. Policymakers must understand that their regulatory actions could constrain development of generative AI and its application across the economy, distorting the market towards closed proprietary models. Thoughtful regulation should aim to protect consumers and society while also fostering an environment where diverse AI models and applications can thrive, thereby ensuring a competitive and dynamic generative AI ecosystem.

Conclusion

We trust that our discussion of frequently asked topics on generative AI – hallucinations, bias, impact on work, and competition and platforms – is helpful for the Committee. Of course, we welcome the opportunity to expand on any of our contributions here by appearing before the Committee at a public hearing.

We look forward to seeing the result of this important inquiry and hearing from you in due course.

Yours faithfully,

Dr Darcy W. E. Allen
RMIT Blockchain Innovation Hub, RMIT University

Professor Chris Berg
RMIT Digital 3 and RMIT Blockchain Innovation Hub, RMIT University

Dr Aaron Lane
RMIT Blockchain Innovation Hub, RMIT University

---

[13] See, for instance, Bartz, D. and Dastin, J. 'OpenAI chief goes before US Congress to propose licenses for building AI', Reuters, 16 May 2023.