

Thank you for the opportunity to respond to the 2019 White Paper *Artificial Intelligence: Governance and Leadership*.

We are broadly in favour of the direction suggested by the White Paper, that is, the establishment of a whole-of-government Responsible Innovation Organisation to oversee the development of AI in Australia.

We write to draw the attention of the Commission to the particular importance and complexity of the uptake of AI *in the health sector in Australia*. Health has lagged in the uptake of AI, but there are now increasingly successful efforts to, for example, apply machine learning and vision to diagnostics, use consumer wearables to monitor and evaluate health states, and provide continuously available chat-bot health coaching. We note that on p. 12 of the White Paper, Accenture's work on the impact of AI on different sectors is mentioned: information and communication, manufacturing and financial services are noted as "the highest annual Gross Value Add growth rates in an AI scenario". Our central point is that the health sector is also a critically important domain in which to regulate and govern AI, for reasons which go beyond market value. This is so because:

1. AI is now developing rapidly in health—we are on the cusp of implementation (Coiera, 2018). This provides a time-limited opportunity to establish good governance prospectively.
2. Health and healthcare carry particular value beyond economic value. Instrumentally, health underpins the ability of individuals to enjoy a life they value (including their ability to participate in market systems). However health also has intrinsic value: to have good health is to have a better life. This increases the importance of good governance of AI in health systems and services.
3. AI will increasingly be integrated into the relationship between healthcare providers and patients, relationships that carry ethical obligations beyond regular market exchange (as recognised in the development of clinical ethics and bioethics). Thus introducing AI into healthcare carries special ethical importance, and requires particular attention. We note also that in as much as there is a recognised human right to access what is necessary for health and wellbeing, including a right to health care (United Nations, 1948), any effect of AI on health systems and services is also a matter for human rights. For this reason, we welcome the AHRC adopting AI as a matter of concern and framing the White Paper in relation to human rights.

Research on the ethics of AI in health is in early development. However it seems likely that the issues that have arisen in other sectors may also be problematic in the context of healthcare and public health, and that these problems are likely to be more salient due to the special value of health to people and the ethical significance of healthcare relationships discussed above. We address these issues in response to relevant questions below.

Consultation questions

Question 1: What should be the main goals of government regulation in the area of artificial intelligence?

In respect of health and healthcare, there are particular concerns that need to be addressed:

1. There is an established and rightful expectation that health service delivery will be *evidence based*. The rules for evidence based healthcare are more stringent than the evidence required for many other forms of service delivery. The social science of biomedicine markets, however, makes clear that enthusiasm about innovation and the need to attract capital often drives 'promissory discourse' – hype without delivery of promised outcomes (Rajan, 2006). One consequence of the enthusiasm for innovation is the uptake into practice of, for example innovative devices, in advance of evidence of their safety and efficacy. Early uptake without evidence of innovative technologies in healthcare can have major impacts on people's health and wellbeing. Recent examples include the failure of the metal on metal hips (Therapeutic Goods Administration, 2017) and more recently, widespread problems caused by use of vaginal mesh (The Senate Community Affairs Reference Committee, 2018). It will be particularly important in healthcare to seek some evidence of beneficial outcomes before AI is widely implemented. We note a particular concern here about the potential for AI systems to be developed (trained and tested) using biased datasets (Committee, 2018).
2. Even within the strictures of evidence based medicine, there are many examples of implementation of services which, counterintuitively, turn out to cause *harm*. This is especially so in screening and early detection, where the problem of *overdiagnosis* is now emerging as an important source of harm (Moynihan et al., 2018). The NHMRC-supported collaboration *Wiser Healthcare*, of which Profs Carter and Houssami are a part, can offer relevant expertise in this regard (Wiser Healthcare, 2017). Many applications of AI in health are in screening and early detection, so have potential to cause overdiagnosis (e.g. AI systems to detect melanoma and breast cancer); the well-recognised general problem in AI of *unintended consequences* is particularly relevant here (Laura, Sreeram, Andrew, & Mustafa, 2017).
3. The maintenance of *trust* in healthcare systems is critical.
4. *Privacy and confidentiality*, always key issues in relation to linked data and AI, are especially ethically significant when they concern health data.
5. Consistent with a human rights framework, the right to conditions for health is a matter for *justice*. It is well-recognised that the structuring of society leads to striking differences in health outcomes for different groups (Commission on Social Determinants of Health, 2008). It seems likely that AI in healthcare may be unequally available to different social groups, and that this could be a further source of health inequity.
6. An authority could work with existing bodies (e.g. specialist medical Colleges, healthcare regulatory authorities) to ensure that AI is implemented in health systems and services consistent with the following principles:

- a. *Explainability* to both healthcare professionals and the general public, consistent with international thinking on AI ethics, to support continued *trust* and *accountability*;
- b. Maximising health equity by preventing any increase in *health inequities*, and actively decreasing inequities (e.g. by making effective care more available in remote areas);
- c. Minimising *harms*, including *unintended consequences*;
- d. Ensuring that the *privacy and confidentiality* of health data are protected;
- e. Supporting regulatory and professional bodies to produce *consistent* rather than piecemeal approaches to interconnected legal, ethical and professional issues such as: responsibility for decision-making, appropriate delegation of decision-making, the potential for rapid de-skilling of professionals, and the problems of risk homeostasis—decreased vigilance when an activity is assumed to be safe—and automation bias, the tendency of humans to over-trust automatic systems and take less personal responsibility (Cretton, 2017);
- f. Ensuring that implementation is supported by *public engagement and deliberative democratic processes* such as community juries (Blacksher, Diebel, Forest, Goold, & Abelson, 2012). The *Australian Centre for Health Engagement, Evidence and Values* at UOW, where Carter is Director, has relevant expertise to support these activities (Degeling, Rychetnik, Street, Thomas, & Carter, 2017).

Question 2: Considering how artificial intelligence is currently regulated and influenced in Australia:

- a. What existing bodies play an important role in this area?

Specialist health professional colleges will need to play a part, but most are not strongly engaged at present. Existing health regulatory bodies including the Australian Commission on Quality and Safety of Healthcare may also be important engagement partners. An Authority could play an important role in prompting engagement on the issue and ensuring consistency across professional groups. In addition the Australian Health Practitioner Regulatory Authority (AHPRA) will need to develop standards for professional behaviour and responsibility in the use of AI as a basis for disciplinary procedures if and when such standards are breached.

- (b) What are the gaps in the current regulatory system?

Question 3: Would there be significant economic and/or social value for Australia in a Responsible Innovation Organisation?

We cannot comment on this overall, but note the importance of health as a domain for AI and our arguments above about the special value of health. If the Organisation was founded and acted in ways that protect human rights

principles and that are transparent and accountable, then this could contribute to maintaining and building trust.

Question 4: Under what circumstances would a Responsible Innovation Organisation add value to your organisation directly?

As academics doing research in this area, it would be useful to have recognised national leadership to guide priority setting.

Question 5. How should the business case for a Responsible Innovation Organisation be measured?

Notwithstanding the political importance of making a business case, we note our arguments above that sometimes actions should be taken for ethical reasons that go beyond financial concerns.

References

- Blacksher, E., Diebel, A., Forest, P.-G., Goold, S. D., & Abelson, J. (2012). What Is Public Deliberation? *Hastings Center Report*, 42(2), 14-17. doi:10.1002/hast.26
- Coiera, E. (2018). The fate of medicine in the time of AI. *Lancet*, 392(10162), 2331-2332. doi:10.1016/s0140-6736(18)31925-1
- Commission on Social Determinants of Health. (2008). *Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health*. Geneva: World Health Organization.
- Committee, S. a. T. (2018). *Algorithms in decision-making: Fourth Report of Session 2017-19*. (HC 351). United Kingdom: House of Commons.
- Cretton, C. (2017). The dangers of AI in health care: risk homeostasis and automation bias. *Towards Data Science*. Retrieved from Towards Data Science website: <https://towardsdatascience.com/the-dangers-of-ai-in-health-care-risk-homeostasis-and-automation-bias-148477a9080f>
- Degeling, C., Rychetnik, L., Street, J., Thomas, R., & Carter, S. M. (2017). Influencing health policy through public deliberation: Lessons learned from two decades of Citizens'/community juries. *Social Science and Medicine*, 179, 166-171. doi:<http://dx.doi.org/10.1016/j.socscimed.2017.03.003>
- Laura, M., Sreeram, V. R., Andrew, P. C., & Mustafa, O. (2017). Unintended consequences of machine learning in medicine? [version 1; referees: 2 approved]. *F1000Research*, Vol 6 (2017). doi:10.12688/f1000research.12693.1

- Moynihan, R., Barratt, A. L., Buchbinder, R., Carter, S. M., Dakin, T., Donovan, J., . . . Scott, I. A. (2018). Australia is responding to the complex challenge of overdiagnosis. *The Medical journal of Australia*, 209(8), 332-334. doi:10.5694/mja17.01138
- Rajan, K. S. (2006). *Biocapital : the constitution of postgenomic life*. Durham: Duke University Press.
- The Senate Community Affairs Reference Committee. (2018). *Number of women in Australia who have had transvaginal mesh implants and related matters*. Canberra: Commonwealth of Australia.
- Therapeutic Goods Administration. (2017). *Metal-on-metal hip replacement implants*. <https://www.tga.gov.au/metal-metal-hip-replacement-implants>: TGA.
- United Nations. (1948). The Universal Declaration of Human Rights. Retrieved from <http://www.un.org/en/documents/udhr/index.shtml>
- Wiser Healthcare. (2017). Wiser Healthcare: A research collaboration for reducing overdiagnosis and overtreatment. Retrieved from <http://wiserhealthcare.org.au/>