

Ian Oppermann

**To cite:** Oppermann I. Regulating AI for health. *BMJ Health Care Inform* 2023;30:e100931. doi:10.1136/bmjhci-2023-100931

Received 10 October 2023  
Accepted 04 December 2023

Healthcare is a unique and complex mix of expert practitioners, small businesses, major providers, professional and semiprofessional contributors. It is highly regulated and procedural, but also an area where ethical issues are regularly tested. It encompasses cutting-edge research and pioneering techniques, as well as large scale applications of well proven treatments and procedures. It is also fair to say there is much about the efficacy of healthcare and treatments that we have only a functional understanding of. After careful trials and review, we know that it does work rather than why it works.

These same characteristics make healthcare a perfect environment for the application of artificial intelligence (AI). The question is how we can use AI, rather than becoming victim to it. Finding ways to think our way into frameworks for regulation and appropriate ways of using AI without it just ‘happening’ to us or possibly worse, ignoring the value it can bring to complex environments such as healthcare.

We have largely been taken by surprise by the tremendous advances in capability of the latest large language models and generative AI. These new data-driven, algorithmic tools have forced us to reconsider the frontier of what we thought AI could do.

In the past, we have used AI to automate, navigate, detect anomalies, recommend next actions, pattern match, predict and explore ‘what if’ simulation scenarios. The newer AI can do that and generate, synthesise, translate and intelligently tackle moderate complexity tasks. They can do this whilst being judged to ‘demonstrate’ greater empathy<sup>1 2</sup> and patience than human respondents in online environments constrained to text-based interactions.

What has changed is the way AI works and the size of the datasets used to train the AI.<sup>3</sup> Generative AI is trained to ‘focus’ and is training on datasets of literally trillions of examples.

This unsupervised training occasionally leads to some surprises. When presented with a supposedly factual response from

your AI query, some results may refer to ‘real world’ sources that simply do not exist. Similarly, a request to generate an image from a verbal description may lead to something a little more ‘Salvador Dali’ like than you may have expected. This scaled up version of an age-old adage of ‘garbage-in-garbage-out’ leads to the modern twist ‘garbage-in-sometimes-hallucination-out’.

So, if we are to use AI, we will need regulation or policies to ensure that it is used appropriately.

### AI IS DIFFERENT TO OTHER TECHNOLOGIES

Some of the concerns raised about AI could just as readily be applied to other technologies when first introduced. If you replaced ‘AI’ with ‘quantum’, ‘laser’, ‘computer’ or even ‘calculator’, some of the same concerns arise about appropriate use, safeguards, fairness, contestability would arise. What is different about AI is that it allows systems, processes and decisions to happen much faster and on a much grander scale. AI is an accelerant and an amplifier. In many cases, it also ‘adapts’, meaning what we design at the beginning is not how it operates over time.

Before developing new rules, existing regulation and policy should be tested to see if it stands up to potential harms and concerns associated with those three ‘a’s’. If your AI also ‘generates’ or synthesises, then a more stress-tests are needed as ‘generation’ goes well beyond what you can expect from your desktop calculator.

### AI IS NO LONGER EXPLAINABLE

Except in the most trivial cases, the depth and complexity of the neural networks (number of layers and number of weights), coupled with the incomprehensibly large training datasets means we have little chance of describing how an output was derived even if it was possible to unpick all of the levels and the impact of each training element. Any explanation would be largely meaningless.



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

University of Technology, Sydney, New South Wales, Australia

#### Correspondence to

Dr Ian Oppermann;  
ian.oppermann@uts.edu.au



For any decision which matters, there must always be an empowered, capable, responsible human in the loop ultimately making that decision. That ‘human-in-the-loop’ cannot just be a rubber stamp extension of the AI-driven process.

### ANY REGULATION MUST NOT REFER TO THE TECHNOLOGY

There have been numerous calls to ban, ‘pause’ or regulate use of AI. The orders of magnitude difference between the pace that technology moves, and that regulation adapts, means the closer the regulation gets to the technology, the sooner it is out of date. Regulation must stay principles based and outcomes focused. Regulation must remain focused on preventing harms, the requirement for appropriate human-based judgement (even if AI assisted), dealing with contestability and remediation.

### WE NEED TO THINK LONG TERM

It is not unreasonable to accept the argument that AI is likely to have as profound an impact as electricity. As AI becomes embedded in devices, tools and systems, it becomes invisible to us. Our expectations of these devices, tools and systems are that they are ‘smarter’: better aligned to the tasks at hand; better able to interpret what we mean rather than what we ask for; and improve over time. We do not expect to be manipulated by, or harmed by the tools we use.

Regulation must provide the oversight to allow us to stay vigilant to any negative consequences from AI use individually, for our society and for the environment. Regulation of AI, especially in healthcare, must be based on safeguards: minimising and addressing harms; and monitoring long-term impacts of use of AI.

In 2020, New South Wales (NSW) developed an AI strategy and AI Ethics policy. In 2021, NSW developed,

tested and mandated the use of an AI Assurance Framework.<sup>4</sup> This framework has strong links to international standards (ISO—the International Standards Organisation—and IEC—the International ElectroTechnical Commission) and is being updated now to accommodate changes in AI capability and is a much faster process than updating regulation.

As our world continues to become more data driven, we will inevitably see more AI used to automate processes, connect systems to identify anomalies in a healthcare context. Our focus must remain on ensuring a safe environment and one which empowers individuals as AI continues to amplify, accelerate and adapt. The genie is out of the bottle, so that focus must also stand the test of time.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

### REFERENCES

- 1 Morrow E, Zidaru T, Ross F, *et al*. Artificial intelligence technologies and compassion in healthcare: a systematic scoping review. *Front Psychol* 2022;13:971044.
- 2 Ayers JW, Poliak A, Dredze M, *et al*. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med* 2023;183:589–96.
- 3 World economic forum. Available: <https://www.weforum.org/agenda/2023/02/generative-ai-explain-algorithms-work/> [Accessed 10 Oct 2023].
- 4 NSW government. Available: <https://www.digital.nsw.gov.au/policy/artificial-intelligence/nsw-artificial-intelligence-assurance-framework> [Accessed 10 Oct 2023].