



Editorial – AI and Safety: Beware of Geeks bearing gifts

New Zealand Journal of Health and Safety Practice; Vol 3, edition 2, July 2026

Author: Robert Birnie, NZISM Risk & Assurance Manager

Email: risk@nzism.org

DOI: <https://doi.org/10.26686/nzjhsp.v3i2.10937>

Introduction

It feels like every day brings a new story about artificial intelligence (AI). Whatever one's opinion on AI, there can be no question that it is a technology that has captured the attention of society.

As NZISM's Risk & Assurance Manager, and a risk professional dealing with technology and AI, I have for some time been considering the promise and perils of AI for health and safety practitioners. One of NZISM's strategic goals is to raise the bar for health and safety practice, increasing the professionalism demonstrated by our members. I would encourage all readers of the Journal to more closely consider how they engage with AI, and technology in general, in service of enhancing their own professionalism.

Included below are several brief thoughts that warrant further consideration, and that I would appreciate your thoughts on.

What's in the Box?

Arthur C. Clarke's Third Law states that "any sufficiently advanced technology is indistinguishable from magic" (Clarke, 1973, p. 39). The Third Law has been applicable to many technologies, for many years, but it is particularly relevant when it comes to AI. A tool that can conjure words, images, sound, even video, from nothing more than a few words? When it works well, it certainly can seem like magic.

Adding to this illusion is the 'black box'. A black box is a system in which the user may observe the inputs and the outputs, but has little, or no idea how such a system works (Ashby, 1956). When the technology of the Third Law is obscured by a black box, it becomes all the more magical.

Is it reasonable to expect a user to understand exactly how the black box works? No.

But we should, at a minimum, understand the potential consequences of relying on a black box. Have you conducted a risk assessment on any black boxes which may be operating in your workplace? What would occur should the black box fail? Would you know it has failed? Would you know if it was ever working?

Retailer Amazon recently announced the closure of their Amazon Go and Amazon Fresh branded stores. These stores offered a simple premise. You take items you want from the shelves and you "Just Walk Out". When you do, your credit card is charged. Magically.

Advertised as using advanced AI, the reality was somewhat less magical. Faced with the reality of the unreliable technology that was supposed to run the stores, Amazon paid thousands of Indian workers to review the transactions of up to 70% of the customers purchasing from the stores (Wayt, 2023).

How can retailers and customers appropriately manage risk when the way in which the black box operates is fundamentally different to that which they were lead to believe?

While this specific example may not have health and safety implications, it evidences the degree of scrutiny that should be applied to any technology, especially one where you cannot observe the inner workings of the machinery.

Buzzword Bingo

AI has become a business buzzword buoyed by an inflating bubble of promises and hype. In reality, the ability for most businesses to effectively execute on an AI project is limited, and the effectiveness

of the end results is questionable. Recent research from MIT found only 5% of AI implementations were delivering value (Challapally et al., 2025).

Challapally et al. specifically highlight the lack of value driven by many adoptions of AI tools, notably large language models (LLMs) such as ChatGPT and Copilot. These broad-brush tools are easy to implement, and straight-forward to use, but what problems are being solved by them? How deeply has their use been thought about?

Like many technology projects, the implementation and usage of AI should be carefully considered. Failing to do so can be a costly, timewasting endeavour. I would encourage you to engage a professional to assist with the implementation of any technology, but if you are unable to do so, please at least consider the following high-level advice:

- Document the business need.
- Define a set of requirements that will meet your business need, and success criteria that will prove these needs have been met.
- Identify and investigate tools until you find one which can provably meet your requirements.
- Plan the acquisition, implementation, and operationalisation of the tool.
- Execute on your plan.
- Measure against your success criteria.

At all stages of your journey, please remember to be inquisitive. Ask questions and challenge the assumptions of yourself and others. The answers you get may surprise you!

Know Thyself

A core tenant of NZISM's Code of Ethics (2022) is that all members must demonstrate a level of professionalism and competence that their position, experience, and qualifications would imply. This requirement is not restricted only to the sphere of occupational health and safety.

The use of advanced technology in the betterment of health and safety outcomes is an admirable goal, but practitioners should consider whether or not they are suitably skilled to provide advice on, support, or even use AI in different circumstances.

I do not say this to outright discourage use of AI tools. But careful consideration of the tool, how you intend to use it, and the purpose you intend to use it for is crucial.

For example, I am not a health and safety practitioner. I have neither the professional experience, nor the hours of study that would give me the confidence to make that claim. As such, it would be foolish of me to ask an AI LLM to provide me with information on health and safety, and then to use this information without a second thought. I would not know if the AI's output is correct, and may lack the knowledge to even be able to determine this without extensive work.

If you are suitably skilled in the content that you are requesting from an AI LLM, you can then determine the validity of the outputs, and apply them to your work appropriately. But please be cognisant of the bounds of your expertise. If you may be operating at the edge of or outside your areas of competence, why not engage a peer who has experience in that field? You may learn a lot!

Finally, I would encourage you to not just know yourself, but value yourself. AI tools can produce useful outputs. But those around you do not engage with you because they want to read the output of an AI LLM.

They want to know what *you* think.

Conclusion

The management of technology-related risk has been part of my professional career for the past 15 years. While the challenges posed by AI are sometimes novel or obfuscated, the management of AI risk is more evolutionary than revolutionary. You do not need to be an expert in the field of AI to engage with it responsibly, you simply need to:

- consider what you know, and what you do not know about the tool, the content, and yourself.

- beware buzzwords and engage in technology-related uplift responsibly; with help from professionals where possible.
- remember your value, and proudly demonstrate it.

NZISM is considering ethical guidelines around the use of AI and want to upskill our members on responsible use. If you have thoughts on the use of AI in occupational health and safety, or if there are AI or technology-related topics you would appreciate NZISM addressing, please do not hesitate to email me: risk@nzism.org.

References

- Ashby, W. R. (1956). *An Introduction to Cybernetics*. London. Chapman & Hall.
<http://pespmc1.vub.ac.be/books/IntroCyb.pdf>
- Challapally, A., Pease, C., Raskar, R., et al. (2025). *The GenAI Divide: State of AI in Business 2025*. MIT Nanda,
https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf
- Clarke, A. C. (1973). Hazards of Prophecy: The Failure of Imagination. In *Profiles of the Future: An Inquiry into the Limits of the Possible*. Pan Books.
https://archive.org/details/profilesoffuture0000clar_q9p1/page/38/mode/2up
- NZISM. (2022). *Code of Ethics*. Retrieved 27/04/2026 from <https://www.nzism.org/code-of-ethics>
- Wayt, T. (2023). *How Amazon's Big Bet on 'Just Walk Out' Stumbled*. Retrieved 27/04/2026 from <https://archive.ph/0iVxq>