



# The breadth of AI risks across Australia; Why we need a considered approach to regulation

Perspectives:  
Susie Alegre, with an  
introduction from  
Reset.Tech

February 2025

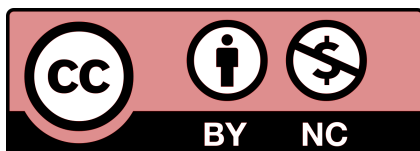
Reset·Tech  
**AUSTRALIA**

## Summary

This paper is a 'thought piece' that offers insights into eight broad areas of concern regarding AI. It is not intended as an anti-AI paper, but rather as a provocation to remind us that as we increasingly adopt AI in our day-to-day lives, there are significant opportunities but also substantial risks. Like all digital risks, these can and should be mitigated with careful regulation.

We hope this paper opens up new areas for regulatory thought as AI regulation takes shape. It offers short provocations addressing:

1. Environmental risks
2. Corporate control of human relationships
3. Creative industries
4. The information ecosystem
5. The rights of women and girls
6. Ultra-processed healthcare
7. Automated injustice
8. Digital dependency



# Contents

Introduction	1
1. Environmental risks	2
2. Corporate Control of “Human” Relationships	3
3. Creative industries	4
4. Information Ecosystem	5
5. The rights of women and girls	6
6. Ultra-processed healthcare	7
7. Automated injustice	8
8. Digital dependency	9

Reset.Tech Australia is an independent, non-partisan policy research lab committed to driving public policy advocacy, research and civic engagement to strengthen our democracy within the context of technology. We are the Australian affiliate of Reset, a global initiative working to counter digital threats to democracy.

## Introduction

While the debate about digital platform regulations continues in Australia – through discussion on social media ‘bans’ for kids, a digital duty of care or ex-ante codes for competition for example – the debates around AI regulation can sometimes feel less prominent. Nevertheless, they continue to progress.

Many of the risks that platform regulation aims to address also exist in the domain of AI. For example; a social media ban will not keep children safe from AI bots that encourage suicide<sup>1</sup>; the ‘content moderation or free speech’ dilemma that plagued the rejected Misinformation and Disinformation Bill reappears in ‘hard wired’ formats as seen in Chinese trained AI censoring responses about Tiananmen Square and Taiwan,<sup>2</sup> and; national security concerns surrounding data-hungry digital platforms<sup>3</sup> are exacerbated by AI training sets.<sup>4</sup> The risks, and opportunities are real.

As discussions about AI regulation emerge in Australia, Reset.Tech Australia wanted to offer a thought piece on the broader contribution about the ‘scope’ of concerns. Just like digital platform regulations, any approach to addressing AI regulation needs to take a broad, risk-based approach. We need to move beyond the narrow concerns that plague platform regulation such as limited takes on ‘safety’ or ‘hacker-focussed’ cybersecurity concerns (as important as they both are) and instead to adopt a systemic focus that sees the breadth of risks and seeks systemic solutions.

We hope this contribution, authored by Susie Alegre, encourages ‘big picture’ thinking, by offering insights into eight often-overlooked areas of concern.

---

<sup>1</sup>Kevin Roose 2024 ‘Can AI be blamed for teen’s suicide?’ *New York Times*

<https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html>

<sup>2</sup>Robert Booth & Dan Milmo 2025 ‘Chinese AI chatbot DeepSeek censors itself in realtime, users report’ *The Guardian*

<https://www.theguardian.com/technology/2025/jan/28/chinese-ai-chatbot-deepseek-censors-itself-in-realtime-users-report>

<sup>3</sup>Irish Council for Civil Liberties 2024 *Australia’s Hidden Security Crisis*

<https://www.iccl.ie/digital-data/australias-hidden-security-crisis/>

<sup>4</sup>Tom Gerkin 2025 ‘Australia bans DeepSeek on government devices over security risk’ *BBC*

<https://www.bbc.co.uk/news/articles/c8d95v0nr1yo>

## 1. Environmental risks

The words we use to talk about AI and emerging technology belie their real world impacts. When we think of the cloud or of virtual worlds, we might be fooled into imagining something quite magical, ethereal and insubstantial. But the reality is much grubbier and poses serious environmental risks.

Data hungry technologies like cloud computing and generative AI, rely on the computer processing power housed in massive data centres worldwide to function. Australia is one of the top five data centre hubs in the world with 214 data centres spread across the country. The data that powers our world requires massive amounts of electricity and water to keep it running.

A single large data centre could use the equivalent energy needed to power 50,000 homes. It is estimated that data centres are currently using 5% of the power on the Australian grid with that figure rising, according to some estimates, to 15% by 2030.<sup>5</sup> This will put pressure on Australia's transition to green energy while also raising serious questions about the impact of increased reliance on AI on climate change.

And AI is also thirsty. Data centres have long been a matter of concern in drought-affected areas because of the water used for cooling. But the shift to generative AI in everything from search to writing a haiku, makes the problem more pressing. Building the tech is thirsty work, but running it is also far more water intensive than the tech we have previously relied on. While a Google search might use half a millilitre of water in the data processing, a short conversation with ChatGPT can use half a litre.

Since the arrival of ChatGPT in late 2022 pushed generative AI into the mainstream of our daily technology use, both water and electricity consumption by major tech companies have spiked. Both Microsoft and Google reported record increases in energy and water consumption in 2023. And Microsoft's energy use has doubled in recent years, rising from 11.2 million megawatt-hours (MWh) in 2020 to 24 million MWh in 2023 while Google saw a 67% increase over the same period. And the carbon footprint of Google alone went up 13% in 2023 and 67% over the past four years. The companies reported increases in water usage for cooling data centres of between 13.8% and 21% in 2023 but this is only a part of the picture – the reporting does not include water use to generate electricity or the risk of water contamination. Projections indicate that, on the current trajectory, 'global demand from AI will be responsible for the use of between 4.2 and 6.6 trillion litres by 2027',<sup>6</sup> equivalent to a third to a half of the current water usage in Australia.<sup>7</sup>

One of the keys to managing the environmental impact of AI is the scale of its use. Using generative AI to search for a holiday or write a haiku is like taking Concorde to the corner shop. Australia needs to balance the risks with the opportunities for a realistic and sustainable adoption of AI.

---

<sup>5</sup>Samuel Yang 2024 'Power-hungry data centres scrambling to find enough electricity to meet demand' Australian Broadcasting Corporation <https://www.abc.net.au/news/2024-07-26/data-centre-electricity-grid-demand/104140808>

<sup>6</sup>Manuel G. Pascual 2024 'Artificial intelligence is already an environmental problem' EL PAIS <https://english.elpais.com/technology/2024-07-16/artificial-intelligence-is-already-an-environmental-problem.html#>

<sup>7</sup>Australia's 2021/2022 water usage in Australia was 13,4149 GL or 13.449 trillion litres. See: Australian Bureau of Statistics 2023 *Water Account Australia* <https://www.abs.gov.au/statistics/environment/environmental-management/water-account-australia/latest-release>

## 2. Corporate Control of “Human” Relationships

In July 2024, a new wearable AI product was launched called ‘Friend’.<sup>8</sup> Marketed as a companion that will always be with you, seeing and responding to the world around you, this is just the latest and most direct example of tech innovation that threatens to render humanity obsolete, even in our own interpersonal relationships. We will now be able to buy a friend for \$99 USD instead of investing the time needed to make a human friend.

‘Friend’ is not alone. The anthropomorphisation of AI to engage and appear to empathise with human emotions is one step further towards corporate capture of our entire lives. Replika, one of the leading providers of AI chatbot relationships, boasts around 30 million users worldwide.<sup>9</sup> Its services cater to an increasingly isolated population and exploit our tech induced vulnerability.

But some Replika users have noticed that their chatbots were displaying coercive and controlling behaviour learned from the interaction with some users who felt liberated to treat their bots in ways that would degrade a real person. Many people choosing chatbot relationships over real-life partners do so in search of a safe space with the perfect companion – one with no baggage. Instead, they find they have all the baggage in the world.

In 2023, the Italian data protection authority found Replika to be illegal in the country, in part because many of its users were children.<sup>10</sup> Shortly after, Replika switched off the capacity for its chatbots to engage in erotic roleplay on its basic service. People who had declared themselves as ‘married’ to their bots found themselves bereft.<sup>11</sup> It was as if a real partner had turned their back on them. Allowing AI companies to sell their products as alternatives to human connection leaves users open to exploitation being asked to pay for an upgrade to keep their intimate lives going. It also makes them vulnerable to manipulation that could harm them or others.

Chatbot relationships have already been implicated in tragic cases of suicide in Belgium<sup>12</sup> and the United States,<sup>13</sup> and an attempt to kill the late Queen in the UK.<sup>14</sup> These virtual relationships are not without real world consequences.

Many users of companion and relationship chatbots are children. This kind of engagement with AI sets unrealistic expectations which may make it harder to form genuine human relationships. Artificial intimacy is being aggressively advertised to young people online, attempting to insert them into communities to build a movement that will break down social ties, leaving Australians lonely and vulnerable. We have to ask whether this is the future we want for our society.

---

<sup>8</sup>Friend n.d. <https://www.friend.com/>

<sup>9</sup>Nilay Patel 2024 ‘Replika CEO Eugenia Kuyda says it’s okay if we end up marrying AI chatbots’ *The Verge* <https://www.theverge.com/24216748/replika-ceo-eugenia-kuyda-ai-companion-chatbots-dating-friendship-decoder-podcast-interview>

<sup>10</sup>Elvira Pollina and Martin Coulter 2023 ‘Italy bans U.S.-based AI chatbot Replika from using personal data’ *Reuters* <https://www.reuters.com/technology/italy-bans-us-based-ai-chatbot-replika-using-personal-data-2023-02-03/>

<sup>11</sup>James Purtill 2023 ‘Replika users fell in love with their AI chatbot companions. Then they lost them’ Australian Broadcasting Corporation

<https://www.abc.net.au/news/science/2023-03-01/replika-users-fell-in-love-with-their-ai-chatbot-companion/102028196>

<sup>12</sup>Chloe Xiang 2023 ‘He Would Still Be Here’: Man Dies by Suicide After Talking with AI Chatbot, Widow Says’ *Vice*

<https://www.vice.com/en/article/man-dies-by-suicide-after-talking-with-ai-chatbot-widow-says/>

<sup>13</sup>Kevin Roose 2024 ‘Can AI be blamed for teen’s suicide?’ *New York Times*

<https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html>

<sup>14</sup>Tom Singleton, Tom Gerken and Liv McMahon 2023 ‘How a chatbot encouraged a man who wanted to kill the Queen’ BBC News <https://www.bbc.co.uk/news/technology-67012224>

### 3. Creative industries

Generative AI is so impressive because it has been trained on the creative output of millions of talented people. The mass theft of copyrighted material from around the world is the business model of this new generation of AI, and the creative industries are bearing the brunt of the damage.

AI companies steal creative work and sell it on as a replacement for human labour. This impoverishes the creative people whose work is stolen, but it also threatens the future of our cultural heritage. AI does not feel, understand, or communicate with us. It has no roots in history or culture, it has no joy or pathos. And the people building and selling it have no sense of the intrinsic value of human creativity or culture.

People working in the creative industries have suffered a wave of despair over the past two years as generative AI has been touted as the future of creative work. Writers find themselves tasked with editing AI generated drivel because their employers have been told it will improve productivity. But there is pushback around the world. Individual writers, artists and publishers have brought lawsuits over copyright breaches<sup>15</sup> and unions brought Hollywood to a standstill in 2023 in an attempt to shore up writers' and actors' rights in the face of AI encroaching on the industry.<sup>16</sup>

Australian copyright law is as yet untested and the outcome of cases in the US will not necessarily be followed here. How the intellectual property, economic and moral rights of creative material are defined will define the future of the cultural and creative space in Australia.

Cultural heritage is a vital part of human society, and the destruction of both tangible and intangible cultural heritage is recognised as an international crime<sup>17</sup> because of its destabilising impact it has on communities. If we allow the creative industries to be swallowed whole by big tech, we may well lose the cultural compass that creativity gives us.

Creativity is vital to the human capacity to reflect, imagine and connect – it is foundational to human civilisation. The creative industries provided work to over 500,000 people in Australia in 2021.<sup>18</sup> Cultural and creative work contributed \$115.2 billion to the Australian economy in 2017 - 2018 alone,<sup>19</sup> but human creativity is ultimately priceless. We cannot afford to lose it.

Countries that protect their creative industries in the face of this threat, by ensuring regulation and the protection of the economic and moral rights of artists and other creatives, will become the incubators of future global cultural life.

---

<sup>15</sup>For example, see: Alexandra Alter and Elizabeth A. Harris 2023 'Franzen, Grisham and Other Prominent Authors Sue OpenAI' *The New York Times* <https://www.nytimes.com/2023/09/20/books/authors-openai-lawsuit-chatgpt-copyright.html>

<sup>16</sup>Andrew Dalton 2023 'AI is the wild card in Hollywood's strikes. Here's an explanation of its unsettling role' *AP News* <https://apnews.com/article/artificial-intelligence-hollywood-strikes-explained-writers-actors-e872bd63ab52c3ea9f7d6e825240a202>

<sup>17</sup>International Criminal Court 2021 *Policy on Cultural Heritage* <https://www.icc-cpi.int/sites/default/files/itemsDocuments/20210614-otp-policy-cultural-heritage-eng.pdf>

<sup>18</sup>Creative Australia 2023 *The Creative Economy in Australia – What Census 2021 Tells Us* <https://creative.gov.au/advocacy-and-research/the-creative-economy-in-australia-what-census-2021-tells-us/>

<sup>19</sup>Department of Infrastructure, Transport, Regional Development, Communications and the Arts 2020 *The economic value of cultural and creative activity* <https://www.infrastructure.gov.au/department/media/news/economic-value-cultural-and-creative-activity-0>

## 4. Information Ecosystem

Journalists, like creative writers, are seeing their livelihoods threatened by generative AI,<sup>20</sup> sealing the fate of independent media already rocked by the financial impact of social media and online news sources. The use of generative AI to produce news articles poses a risk to the entire information ecosystem, already polluted by online misinformation and disinformation.

Generative AI is not a research tool, it is a probability machine. If you ask it a question it will respond with the most likely sequence of words based on its training data set. The outputs have nothing to do with the truth, and as AI is increasingly trained on the reams of synthetic AI generated content flooding the internet, the risks of incoherence, bias, and ultimately model collapse, only grow as AI effectively eats itself.

AI cannot find a news story, identify and engage with sources or corroborate and analyse anything that has not been done before. Without trusted sources of information, we don't know what to think, policymaking becomes a lottery, and communities become divided. In Australia, and other places around the world, cases are already cropping up of people suing for defamation by generative AI with the output of AI models representing them in ways that destroy their reputation.<sup>21</sup> If AI driven search engines become the norm, our ability to find the truth at all online will be lost entirely with AI hallucinations and deliberately created fake news providing corroboration for dangerous untruths.

A free and independent media is a cornerstone of a democratic society. Without trusted sources of information, we don't know what to think. Protecting the media from the encroachment of generative AI is vital to protect free speech and freedom of information.

Academic research has also been compromised by the twin drivers of greater productivity and technological innovation, which are often prioritised over quality and integrity. While students' work is increasingly scrutinised by tech solutions to the tech problem of generative AI in academia, scholars are also increasing their research outputs by publishing AI generated nonsense in academic journals.

In Australia, the science magazine Cosmos has drawn criticism for its decision to publish AI generated articles after hitting financial troubles and losing half its staff in 2024.<sup>22</sup> The academic publisher Wiley shut 19 of its scientific journals last year<sup>23</sup> due to their content being flooded with meaningless AI generated materials, highlighting deeper systemic problems in the academic research publishing model. Widespread use of AI in scholarship is a threat to science and to the human capacity for critical thinking.

An already fragile information ecosystem undermined by big business models, AI supercharges the problem by polluting at source.

---

<sup>20</sup>For example, see: Reuters 2024 *OpenAI to use FT content for training AI models in latest media tie-up* <https://www.reuters.com/technology/financial-times-openai-sign-content-licensing-partnership-2024-04-29/>

<sup>21</sup>Anastasia Santoreneos 2023 'Australian politician sues ChatGPT for defamation in landmark case' *Forbes Australia* <https://www.forbes.com.au/news/innovation/australian-politician-sues-chatgpt-defamation-landmark-case/>

<sup>22</sup>James Purtill 2024 'Cosmos Magazine publishes AI-generated articles, drawing criticism from journalists, co-founders' *Australian Broadcasting Corporation*

<https://www.abc.net.au/news/science/2024-08-08/csiro-cosmos-magazine-generating-articles-using-ai/104186330>

<sup>23</sup>Thomas Claburn 2024 'Wiley shuts 19 scholarly journals amid AI paper mill problems' *The Register* [https://www.theregister.com/2024/05/16/wiley\\_journals\\_ai/](https://www.theregister.com/2024/05/16/wiley_journals_ai/)



## 5. The rights of women and girls

AI chatbots that sound eerily like Scarlett Johansson (without her consent)<sup>24</sup> and gynoid robots like Sophia the Saudi citizen robot,<sup>25</sup> may appear to be the friendly face of AI, but the anthropomorphisation of AI through the female form has serious implications for women's rights.

The trend in feminine voices for AI assistants can entrench gender stereotypes of women as willing servants. But the advent of AI CEOs with female names and avatars does not enhance women's representation on boards – it glosses over the problem, replacing women with technology.

In 2023, the DevTernity conference was cancelled following a backlash when it was discovered that its advertised female speakers from the tech world were fake AI generated profiles created to conceal the lack of diversity in the lineup.<sup>26</sup> Meanwhile, AI generated supermodels like Shudu presented as a Black South African woman, make fortunes for their white male creators<sup>27</sup> while displacing real women from the workplace.

Perhaps as predictably as AI text, AI representations of women are often derogatory. The spread of misogyny and pornography online has been supercharged by the speed of mass creation afforded by easy access to generative AI tools. Just last year, Taylor Swift shone a celebrity spotlight on the threat of deepfake pornography or nonconsensual image based sexual abuse. Unsurprisingly, perhaps, in 2023, 98% percent of deepfake videos online fell into this category and 99% of those featured women.<sup>28</sup>

The widespread availability of AI tools and apps like 'Nudify' that allow users to create naked images of anyone from a photo, means that celebrities are no longer the only targets of this. In Australia school girls have been victims of this kind of image abuse created and shared by their teenage peers.<sup>29</sup> Women with a public profile, like politicians, have been targeted as well.<sup>30</sup> But any woman is a potential target. In 2024, Andrew Hayler, a bartender in New South Wales was sentenced for creating and posting hateful content and degrading doctored images of his female friends and colleagues on online pornographic websites.<sup>31</sup>

The impact of non-consensual image based sexual abuse on women and girls is incalculable. It degrades those targeted and puts them in both physical and psychological danger. It can be used personally and politically to undermine and demean the target. AI, when used in these ways, poses a massive threat to the dignity and rights of all women and girls in Australia.

---

<sup>24</sup>Tripp Mickle 2024 'Scarlett Johansson Said No, but OpenAI's Virtual Assistant Sounds Just Like Her' *The New York Times* <https://www.nytimes.com/2024/05/20/technology/scarlett-johansson-openai-voice.html>

<sup>25</sup>Wikipedia n.d. 'Sophie' *Wikipedia* [https://en.wikipedia.org/wiki/Sophia\\_\(robot\)](https://en.wikipedia.org/wiki/Sophia_(robot))

<sup>26</sup>Natalie Lung and Ella Ceron 2023 'Developer Conference Axed After Fake Female Profiles Outcry' *Bloomberg* <https://www.bloomberg.com/news/articles/2023-11-28/tech-conference-faces-backlash-on-claims-of-fake-women-speakers>

<sup>27</sup>Sara Semic 2019 'Meet The Man Behind The World's First Digital Supermodel' *ELLE* <https://www.elle.com/uk/fashion/a28394357/man-behind-worlds-first-digital-supermodel/>

<sup>28</sup>Security Hero 2023 '2023 State of Deepfakes' <https://www.securityhero.io/state-of-deepfakes/#key-findings>

<sup>29</sup>Angus Watson and Hilary Whiteman 2024 'Teenager questioned after explicit AI deepfakes of dozens of schoolgirls shared online' *CNN* <https://edition.cnn.com/2024/06/13/australia/australia-boy-arrested-deepfakes-schoolgirls-intl-hnk/>

<sup>30</sup>Laura Lavelle 2023 'Antonio Rotondo guilty of contempt of court after allegedly creating deepfake images of school students and teachers' *Australian Broadcasting Corporation*

<https://www.abc.net.au/news/2023-12-06/qld-deepfake-images-court-charge-antonio-rotondo-school-students/103195578>

<sup>31</sup>Ruby Cornish 2024 'Sydney bartender Andrew Hayler jailed after sharing digitally altered images of women on porn site' *Australian Broadcasting Corporation* <https://www.abc.net.au/news/2024-06-21/nsw-bartender-jailed-sharing-fake-images-women-on-porn-site/104005942>

## 6. Ultra-processed healthcare

Health is one of the areas where AI is billed as being transformative offering the opportunity for huge leaps forward in human knowledge.<sup>32</sup> AI drug discovery and precision medicine promise to eradicate diseases and find a cure for cancer.

AI will undoubtedly play an important role in the future of medicine, but AI is many things and the sales pitch for AI in medicine is often filled with smoke and mirrors that obscure the risks and undermine the benefits. The AI that helps radiologists identify subtle changes in patients' cells to detect disease early on, is not the same AI that will write you a personalised prescription. There is a risk that useful medical developments will be subsumed by AI hype.

In recent years, overstatement of the potential of innovative Med-tech, has resulted in funds being diverted into useless devices and patients being put in danger.

Theranos, a company that promised near-magical diagnosis through blood tests from a pinprick drop of blood, received \$700 million USD of investment and was lauded as the future of healthcare with a value of \$10 billion USD at its peak. That was before its founders were convicted of fraud, their tech proving to be a dangerous mirage.<sup>33</sup>

Calmara, an app that promised to identify sexually transmitted diseases with 94.4% accuracy almost instantly through AI analysis of a photo of a potential partners' penis,<sup>34</sup> was shut down by the Federal Trade Commission in the US within six months of its launch because there was no evidence to support the medical claims.<sup>35</sup> It is unclear how many people may have contracted STDs after relying on the dubious technology before it was closed down. Excessive faith in the potential of AI in medicine is a serious risk to individuals and community health.

AI as a solution to the costs of access to mental health support is another area of risk. AI therapists are already available for free online, offering 24-hour access without judgement to help you through mental health issues.<sup>36</sup> These services fail to understand the serious needs and potential risks for patients seeking mental health support and risk exacerbating conditions in ways that pose a danger to users and those around them.

In order to harness the potential of AI and medicine, we need to recognise the risks of AI snake oil and regulate to eradicate them.

---

<sup>32</sup>Junaid Bajwa et al. 2021 'Artificial intelligence in healthcare: transforming the practice of medicine' *Future healthcare journal* vol. 8,2 (2021): e188-e194. doi:10.7861/fhj.2021-0095 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8285156/>

<sup>33</sup>Bernd Debusmann Jr & James Clayton 2020 'Theranos CEO Elizabeth Holmes begins 11-year prison sentence' *BBC News* <https://www.bbc.com/news/world-us-canada-65756588>

<sup>34</sup>Corinne Purtill 2024 'An AI app claims it can detect sexually transmitted infections. Doctors say it's a disaster' *Los Angeles Times* <https://www.latimes.com/science/story/2024-04-07/calmara-claims-it-can-detect-stis-doctors-say-its-a-disaster>

<sup>35</sup>Federal Trade Commission 2024 *HeHealth/Calmara* <https://www.ftc.gov/legal-library/browse/cases-proceedings/closing-letters/hehealthcalmara>

<sup>36</sup>Financial Times 2023 *Mental health apps: the AI therapist cannot see you now* <https://www.ft.com/content/e0730064-7b24-4556-9322-45a806c4c5f7>

## 7. Automated injustice

Law tech is another area where AI is being pushed as a way of boosting productivity and lowering legal costs. Lawyers and judges around the world have been experimenting with generative AI as a tool to synthesise the law, search through and reproduce contracts or improve access to justice without too much human effort.<sup>37</sup> But while generative AI can deliver coherent text with apparent confidence, a significant risk is that, while plausible, it is often wrong.

AI is not a truth machine. It does not understand the data it is trained on, nor does it have professional ethics. It is a tech-enabled bullshitter. Lawyers have been caught out making submissions in court with entirely synthetic case law,<sup>38</sup> defendants presenting meaningless AI generated character references,<sup>39</sup> and some tech companies have met challenges when offering robot lawyers who lack actual legal qualifications to back up their services.<sup>40</sup>

Research comparing the outputs of LLMs providing legal analysis trained on genuine judgments, found them to be incorrect at least 58% of the time<sup>41</sup> – a failure rate that would likely see a human lawyer disbarred, or at least facing questions of liability. Less tailored models have even less chance of producing correct assessments of the law.

Judicial decisions based on AI, risk being both inaccurate and unjust. The way LLMs function does not allow for reasoning in the way we expect from a reasoned judgement.<sup>42</sup> The outcome may appear plausible but it is impossible to identify the real reasons for it. AI is already used in sentencing or parole decisions<sup>43</sup> by algorithms parsing data to decide on risks for offending. But because the reasons for risk ranking are unclear, it is difficult to mitigate risks or to ensure that outcomes are not discriminatory based on historical biases in the data.<sup>44</sup>

Increased automation in the justice system also runs the risk of automation bias - our tendency to believe computer generated information over the evidence in front of our noses – or automation complacency, our inclination not to check a computer's work.

In the UK, a glitch in accounting software at the Post Office, resulted in the wrongful conviction of hundreds of subpostmasters over 20 years for accounting crimes and fraud that quite simply never happened.<sup>45</sup> When the figures didn't add up, it was the users, not the tech that were blamed, resulting in the most widespread miscarriage of justice in British history.

By automating justice, we risk automating injustice on a grand scale.

---

<sup>37</sup>Claudia Williams 2024 'Want to use ChatGPT to help prepare for court? This is what lawyers say you should and shouldn't do' *Australian Broadcasting Corporation*

<https://www.abc.net.au/news/2024-05-23/generative-ai-chatbots-responsible-use-in-court-guidelines/103863968>

<sup>38</sup>Damien Carrick and Sophie Kesteven 2023 'This US lawyer used ChatGPT to research a legal brief with embarrassing results. We could all learn from his error' *Australian Broadcasting Corporation*

<https://www.abc.net.au/news/2023-06-24/us-lawyer-uses-chatgpt-to-research-case-with-embarrassing-result/102490068>

<sup>39</sup>Henry Ea 2024 'ChatGPT Infiltrates Australian Courtroom' *LSJ*

<https://lsj.com.au/articles/chatgpt-infiltrates-australian-courtroom/>

<sup>40</sup>Bobby Allyn 2023 'A robot was scheduled to argue in court, then came the jail threats' *NPR*

<https://www.npr.org/2023/01/25/1151435033/a-robot-was-scheduled-to-argue-in-court-then-came-the-jail-threats>

<sup>41</sup>Matthew Dahl, Varun Magesh, Mirac Suzgun, Daniel E Ho 2024 'Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models' *Journal of Legal Analysis*, Volume 16, Issue 1, 2024, Pages 64–93, <https://doi.org/10.1093/jla/laae003>

<sup>42</sup>Perkin Amalaraj 2024 'Decision by judge to use ChatGPT to reach a verdict sparks fury' *Daily Mail*

<https://www.msn.com/en-sg/news/other/decision-by-judge-to-use-chatgpt-to-reach-a-verdict-sparks-fury/ar-AA1ogb8C>

<sup>43</sup>Karen Hao 2019 'AI is sending people to jail—and getting it wrong' *MIT Technology Review*

<https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/>

<sup>44</sup>Aleš Završnik 2019 'Algorithmic justice: Algorithms and big data in criminal justice settings' *European Journal of Criminology*, 18(5), 623-642. <https://doi.org/10.1177/1477370819876762>

<sup>45</sup>BBC 2024 *Post Office Horizon scandal: Why hundreds were wrongly prosecuted*

<https://www.bbc.co.uk/news/business-56718036>

## 8. Digital dependency

The pressure on businesses and the public sector to boost productivity, lower costs and improve access by harnessing the power of AI risks creates widespread digital dependency without serious consideration of the consequences. Scared by the threat of being left behind or becoming obsolete, businesses and governments are rushing to insert AI into places it should never be and building reliance on inherently unreliable and untested systems.

Research into the intensive use of GPS systems suggests that relying on GPS actually reduces our innate capacity to find our own way around<sup>46</sup> without it. Similarly, studies on generative AI show initial signs that while students score higher while using the technology, their scores drop below their initial level when the tech is taken away.<sup>47</sup> Reliance on AI tools for cognitive tasks may, in the long run, undermine our capacity for critical thinking in ways we cannot yet predict.

The global IT outage that started with an update to security software in Australia in July 2024 grounded planes, disrupted banking and medical operations and closed businesses around the world.<sup>48</sup> The cost to the Australian and the global economy remains to be quantified. Though the glitch was quickly identified, it should serve as a wake-up call about the fragility of vital infrastructure that relies on AI and other technology. This is particularly vulnerable when kept in the hands of a small number of global companies.

If a tech update can cause such chaos worldwide, imagine the impact of a cyberattack or a power outage on basic public services. Hospitals,<sup>49</sup> defence facilities,<sup>50</sup> libraries<sup>51</sup> and transport networks<sup>52</sup> have been hit by cyberattacks from criminals or hostile states in Australia and around the world. The increase in extreme weather events adds to the risks of fire or flood taking out essential infrastructure or destroying vital data in ways we are unprepared for.<sup>53</sup>

Perhaps the biggest risk of AI to Australians is the risk of digital dependency. In adopting AI, we need to think about tech security and the risks of monopolisation, but also we must be prepared to fall back on human skills when the lights go out.

---

<sup>46</sup>Louisa Dahmani & Véronique D. Bohbot 2020 'Habitual use of GPS negatively impacts spatial memory during self-guided navigation' *Sci Rep* 10, 6310 (2020) <https://doi.org/10.1038/s41598-020-62877-0>

<sup>47</sup>Muhammad Abbas, Farooq Ahmed Jam and Tariq Iqbal Khan 2024 'Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students' *International Journal of Educational Technology in Higher Education* 21, 10 (2024) <https://doi.org/10.1186/s41239-024-00444-7>

<sup>48</sup>Annika Burgess, et al. 2024 'Australians have been hit by a worldwide tech outage. Here's what we know' *Australian Broadcasting Corporation* <https://www.abc.net.au/news/2024-07-19/global-tech-outage-latest/104120106>

<sup>49</sup>Benita Kolovos 2023 'St Vincent's Health Australia says data stolen in cyber-attack' *The Guardian* <https://www.theguardian.com/australia-news/2023/dec/22/st-vincent-s-health-australia-hack-cyberattack-data-stolen-hospital-aged-care-what-to-do>

<sup>50</sup>Ty Roush 2023 'Russian Hackers Breached 632,000 DOJ And Pentagon Email Addresses In Massive MOVEit Cyberattack, Report Says' *Forbes* <https://www.forbes.com/sites/tyleroush/2023/10/30/russian-hackers-breached-632000-doj-and-pentagon-email-addresses-in-massive-moveit-cyberattack-report-says/>

<sup>51</sup>British Library cyber incident review 2024 *Learning Lessons from the Cyber Attack* <https://www.bl.uk/home/british-library-cyber-incident-review-8-march-2024.pdf/>

<sup>52</sup>Nick Huber 2022 'Transport systems give hackers a moving target' *Financial Times* <https://www.ft.com/content/dc3cccb0-533c-4c5b-bc20-5ee3d422909f>

<sup>53</sup>DCS 2022 *How are Data Centers Affected by Extreme Weather?* <https://blog.datacentersystems.com/how-are-data-centers-affected-by-extreme-weather>