Journal & Proceedings of the Royal Society of New South Wales, vol. 158, part 1, 2025, pp. 119–129. ISSN 0035-9173/25/010119-11

2024 Royal Society of New South Wales and Learned Academies Forum: "Threats to Democracy"

Panel Session 3: Technological challenges to democracy[1]

Prof. Sally Cripps FRSN (Chair), Director, Technology, Human Technology Inst., UTS
sally.cripps@uts.edu.au

Prof. Ed Santow FAAL Inst., UTS, Director, Policy & Governance, Human Technology Inst, UTS
Edward.Santow@uts.edu.au

Dr Darren Saunders, NSW Deputy Chief Scientist & Engineer
Darren.saunders@chiefscientist.nsw.gov.au

Assoc Prof Fatemeh Vafaee, Sch. Of Biotechnology & Biomolecular Science, UNSW
f.vafaee@unsw.edu.au

Sally Cripps: My name is Sally Cripps, and this session is titled *Threats? to Democracy*. The question mark is important — I'd also like us to explore not only the threats, but the potential opportunities that technology can offer.

To set the scene before introducing our fantastic speakers, I want to talk briefly about the relationship between democracy and technology over the last 200 years. There's a graph I often use that shows the various forms of democracy and how they've changed from around 1780 up to the present. Of course, this evolution didn't happen in isolation — many things were happening in the world during that time.

I'm a statistician, and I don't know if any of you are familiar with the wonderful Hans Rosling. He might just have been the world's only charismatic statistician. He created an animated presentation for the BBC, showing in real time how the health and wealth of societies changed from 1810 through to 1960, and then up to 2009 — which was when the presentation was made.[2] His key message was that back in 1810, almost everyone was poor and sick. Life expectancy was under 40 years. But over time, aligned with the rise of democracy, we saw an enormous expansion in both health and wealth across the globe. Rosling attributes this to the rise of technology — particularly the Industrial Revolution — and when you look at the data, it's a compelling story.

However, when you plot technology against democracy directly — for example, industrial production versus democratic development — you see that the relationship is not linear. It's complex and has shifted over time. In the so-called "Swinging 60s," for example, we saw a particular phase of development.

The worrying part, though, is what's happened in the past decade. On the democracy graph, after decades of growth, we see a clear decline. That's what has been alarming many people — those data points are now much

1 This is an edited transcript of the session, which can be viewed at https://www.youtube.com/watch?v=MvP5X5_5i6Y

2 https://www.youtube.com/watch?v=jbkSRLYSojo

lower than what we might expect given the continued advancement of technology.

So, what's going on in the relationship between technology and democracy? How can we better understand the present moment, both in historical context and looking ahead? Let me put forward a few ideas.

First, and perhaps unsurprisingly, technologies like ChatGPT and generative AI have enabled misinformation and disinformation at an unprecedented scale. My brilliant colleagues at the Human Technology Institute have written extensively about facial recognition technology and the way it invisibly violates privacy — again, at a speed and scale we've never seen before. And these issues — privacy, access to reliable information — are fundamental to functioning democracies, as our speakers this morning reminded us.

But there is also an upside. I had the privilege of meeting Audrey Tang, Taiwan's first Digital Minister. She made remarkable use of digital platforms to enhance democracy. In fact, under her leadership, the government's approval rating rose from just 9% in 2014 to 90% in 2020 — it's since dropped slightly to 60%, but still, that's extraordinary. In Taiwan's case, digital platforms helped build trust between the government and the people.

Hopefully, that gives you a picture of how technology can sometimes support democracy and sometimes undermine it.

Now, it's my pleasure to introduce our speakers, who will explore these issues from different perspectives.

Our first speaker is Professor Ed Santow. Ed is the co-founder and co-director of the Human Technology Institute at UTS. He's also a former Australian Human Rights Commissioner, a board member of several charities, and — along with Professor Nicholas Davis — one of the country's leading experts on AI regulation and governance. Ed will speak about the impact of AI on human rights, drawing on his excellent recent book.[3]

Our second speaker is Dr Darren Saunders, the Deputy Chief Scientist of New South Wales. Darren has spent more than two decades working across academia and industry, with a background in science, biology, and neurology. He's also an outstanding communicator and advocate for making science accessible to the public.

Our final speaker is Associate Professor Fatemeh Vafaee. She is based at the School of Biotechnology and Biomolecular Science at the University of New South Wales and serves as Deputy Director of the Data Science Hub. Her work focuses on applying AI to medical contexts, right down to the cellular level. Today, she'll talk about the potential benefits of that work — and also the risks it could pose, and what harm might look like in practice. Without further ado, I'll hand over to Ed.

Ed Santow

I'm going to talk about the intersection of freedom of expression, technology — particularly artificial intelligence — and our democracy. I say that because I'm going to take a somewhat circular route to get there, but rest assured: I will reach that destination.

Let me begin with a name that I suspect few, if any of you, have heard: Stephen Ayres. If you were of a more writerly bent and tried to sketch the true 21st-century

3 Santow E and Mellor D (2024) *Machines in Our Image: The Need for Human Rights in the Age of AI*. LexisNexis.

American everyman, it might look a lot like him. Ayres spent most of his adult life in employment — not wealthy, not politically engaged—but deeply proud of his American identity. Ayres will be remembered, if at all, because he committed a serious crime. He was one of those who stormed the U.S. Capitol on January 6, 2021. He was convicted for offences connected to what can reasonably be described as an attempted *coup d'état*.

What's interesting is his own reflection on how he ended up there. He has spoken candidly about his motivations. Three things stand out.

First, he had disengaged from conventional, authoritative sources of news and information. Second, his worldview was shaped by a deep sense of nationalism and a perception that others were progressing more quickly than he was. Third, and perhaps most importantly, his understanding of the world came almost exclusively from social media — Facebook and what was then called Twitter.

From that extremely narrow information diet, he became utterly convinced that the 2020 election had been stolen from Donald Trump. And he believed, as a patriotic American, that the only proper course of action was to act. Now, I suspect very few in this room would share that worldview. But it's important to try to understand how someone could arrive there.

1. Freedom of expression

This leads me to the first of three key points: freedom of expression.

We often hear — from people like me, or from Emeritus Professor Rosalind Croucher — that freedom of expression is one of the critical preconditions of a healthy, functioning democracy. And that's true. But there's a catch: too often, we think of free expression only in terms of the right to speak — the right to be "on transmit," so to speak.

Certainly, we have never lived in an era where that right has been more accessible. Despite the complaints of certain billionaires, we all now hold a virtual microphone. Social media gives everyone a platform to speak.

But freedom of expression has two critical elements: not just the right to speak, but also the right to *receive* information. If the information you receive is so polluted — by falsehoods, ideological distortion, and disinformation — that you can no longer form a rational view of the world around you, then you are not truly free to express your views either. Because your thinking itself has been manipulated. And that, I think, is one of the major challenges we now face in sustaining a functioning democracy.

2. Social media and new technologies like AI

This brings me to my second point: social media and new technologies, particularly artificial intelligence.

For a long time, people like me — rather embarrassingly — have said that we're living through a dramatic rise in the volume of hoaxes and false information. It felt intuitively true. But recent research[4] shows something surprising: there hasn't been a

4 Uscinski J et al. (2020) Have beliefs in conspiracy theories increased over time? *PloS one* 17(7);
Osman M (2023) Conspiracy theories aren't on the rise — we need to stop panicking. *The Conversation*, 20 June;
Park S et al. (2020) Global mistrust in news: the impact of social media on trust. *International Journal on Media Management* 22(2): 83–96.

significant increase in the *amount* of false information circulating.

Instead, what we're experiencing is something more subtle and more dangerous: a *collapse in the authority of truth*.

What I mean by that is this: in previous decades, people like Chief Scientists, public health officials, or experienced journalists could say something clear — like "Don't inject disinfectant to cure COVID" — and it would be believed. Not by everyone, but by a critical mass. Their statements had authority. There was a common reference point for facts. Now, that's changed. Truth and falsehood are increasingly treated as morally or politically equivalent. That's been hastened by the platforms we use.

3. The attention economy

This leads me to my third and final point: what's accelerating this collapse?

There are many factors, but one of the biggest is the design of social media platforms themselves. When people once got their news from regulated, professional media organisations — however imperfect those organisations were — there were checks and balances. Journalists operated in a market where truth mattered. If they consistently published falsehoods, they would suffer reputational and commercial damage. Regulators and editorial oversight — however flawed — still mattered.

Social media is different. These platforms are not regulated like media companies, and their currency isn't truth. It's attention. The "attention economy" rewards content that keeps you looking longer — whether it's true or not.

The algorithms that drive social media don't have a political agenda, but they have shown us one undeniable thing: the best way to hold someone's attention is not with calm, well-evidenced statements from human rights commissioners or scientists. It's with extreme views, emotionally charged content, and polarising narratives.

Conclusion

Put all that together, and, yes — I am worried about democracy. I share the perspective of Jeni Whalan and Nick Bryant, who spoke earlier today: we're not at a point of collapse in Australia. But we are at a point of serious risk.

We need to push back on three fronts:

1. *Freedom of expression* must be understood as both speaking and receiving reliable information.
2. *Our information environment* cannot be dominated by platforms that treat truth and falsehood equally.
3. *Social media platforms* must be properly regulated — not only to moderate harm, but to safeguard democratic functioning.

If we can address those challenges, we will be in a better position to protect and sustain our democracy. Thank you.

SC: Thank you, Ed. That was absolutely wonderful. I love things in threes — and the point about misinformation not increasing, but our declining ability to debunk it, is absolutely fascinating. Now it's my very great pleasure to introduce Darren Saunders.

Darren Saunders

Now, I want to very quickly talk about the tension between the risks and benefits of technology. I'll focus on a few examples from my own field, because one of the key points I want to make is this: the threat to democracy doesn't necessarily come from

the technology itself, but from how it's applied and perceived.

There are some unbelievably powerful technologies currently reshaping not just everyday life, but also how scientists like me understand the natural world — particularly the human body and brain.

Take fruit flies, for instance. They're a favourite tool for geneticists. A fruit fly brain contains about 140,000 neurones — the human brain, by comparison, contains tens of billions. Recently, researchers created a complete wiring diagram of the fruit fly brain. They sliced a brain into 7,000 sections, ran each through an electron microscope to produce 21 million images, and then used AI to reassemble it into a 3D map. This wiring diagram lets you trace how a taste of sugar activates specific neurones and triggers muscle movement. It's a small but profound example of how technology is transforming fundamental biological research.

We're now trying to do the same with the human brain — an even more complex challenge. Here's another example: if you look up into the Milky Way and see billions of stars, that's roughly the same number of protein molecules in a single brain cell — and there are billions of those cells in every human brain. It's an almost impossible problem to understand on a human timescale without advanced technology. That's why AI has been so revolutionary, particularly in the field of protein folding. The 2023 Nobel Prize in Chemistry was awarded for work in this area. Why does protein folding matter? Because misfolded proteins cause diseases like Alzheimer's and motor neurone disease. Understanding how they fold is a game-changer.

Another area is synthetic biology — where we manipulate genomes to create entirely new forms of life. This has major applications: in agriculture, to create new food, fuel, and fibre; in medicine; in decarbonization; and even in semiconductor design. These engineered life forms are not only philosophically new, but fall completely outside current regulatory frameworks.

Let me share one more example: Athena — an AI tool recently adopted by the NSW Rural Fire Service. Athena aggregates data from weather satellites, on-the-ground reports, aircraft surveillance, and even social media — like geotagged images of fire trucks or smoke. It combines this with CSIRO's decades of modelling to predict the impact of fires and help allocate resources. This is another powerful application of technology that most people never think about, but that saves lives.

Now let's look at genomics — a field where some of these challenges are most evident. You've probably heard of the Human Genome Project. But as genomics becomes more embedded in our healthcare system, it raises new ethical and social issues — especially around identity, privacy, and control.

One high-profile example is the company 23andMe. For a few hundred dollars, you could send in a saliva sample and get a detailed genetic profile — your health risks, ancestry, and more. But what most users didn't realise was that the company's business model relied on collecting and selling that data to pharmaceutical companies. That company recently filed for bankruptcy, and may be sold to another entity.[5]

Now millions of users are worried: "Who owns my data? What are they doing with it?"

5 In March 2025, 23andMe filed for Chapter 11 bankruptcy. [Ed.]

The implications are vast — for insurance, family histories, and personal privacy. These issues go right to the heart of democratic control.

Let's return to the notion of control and misinformation — two themes Ed touched on. In medicine, genomic technology has created an expectation that your genome will be decoded by a doctor who then hands you a personalised treatment. That's the dream. And for a few rare conditions, it's a reality. For instance, spinal muscular atrophy — a fatal condition affecting children — now has a genetic therapy that is literally saving lives. But it costs millions of dollars per patient. It raises huge questions around cost, access, and fairness.

The broader problem is a mismatch between hype and reality. Most people won't receive a personalised treatment when they walk into a hospital. That gap breeds disappointment and mistrust — and it's exactly the space exploited by misinformation peddlers like Pete Evans and Belle Gibson.[6] These influencers profited off people's frustrations, offering false hope in place of scientific medicine.

This all came to a head during COVID-19. Trust in public health, science, and technology was badly shaken. People conflated frustrations over lockdowns and vaccine mandates with distrust in the science itself. That's when we started seeing truly bizarre theories — like 5G networks controlling people through vaccines. It's worth noting the irony that most conspiracy theories were spread via the very same 5G-connected smartphones.

And here's the kicker: even when scientists like me went on TV to debunk these myths — like Donald Trump's suggestion to inject disinfectant — we may have inadvertently amplified the misinformation. There's solid evidence that even addressing false claims can reinforce them in people's minds. It's a paradox. A wicked, unsolvable problem.

So, what are the core challenges and risks?

- *Equity*: in medicine, and beyond, we face serious inequities in how technology is accessed and applied. Without fair access, we undermine trust — and, as Ed said, that's a big problem for democracy.
- *Bias and assumptions*: many datasets and models are built using people who look like me — white, male, Western — and that excludes much of the world's diversity. That skews outcomes, and it's dangerous.
- *Control*: who controls the tech and the data? That's where trust often breaks down.
- *Surveillance*: a few years ago, researchers swabbed subway handles in New York and sequenced the DNA. They not only found traces of the plague in rats, but could also identify the ethnic profiles of entire neighbourhoods. That level of biological surveillance raises deep ethical concerns.
- *Misinformation and misunderstanding*: we're already seeing this play out. Genetically modified food. COVID vaccines. And it's likely to get worse as new vaccines developed with advanced technologies hit the market. Hesitancy and mistrust will follow if we're not prepared.
- *Technology, evidence, and policy*: often, what the technology tells us — "Here's the problem, here's the fix" — doesn't align with how policy works. That mis-

6 See Netflix TV series "Apple Cider Vinegar." [Ed.]

alignment creates disillusionment and makes people feel shut out of democratic decision-making.

So I'll stop there. Hopefully, that gives you a few points for discussion.

SC: Thank you, Darren. It's amazing to hear the breadth of your work and insights. You drew an important connection between equity and democracy. A fantastic talk — thank you. Now, speaking of equitable access to health, I'd like to welcome our final speaker, Fatemeh.

Fatemeh Vafaee

AI opportunities to enhance democratic principles

I see myself as an "AI citizen" — I've been in this field for over 15 years. I completed my PhD in computer science and artificial intelligence back in 2011, right when deep learning models were first being published. Since then, I've contributed to the field through research, leadership in biomedicine, and entrepreneurship. I'm an associate professor, a team leader, and I run a proprietary company focused on translating AI innovations into practical healthcare solutions that directly reach patients.

Today, I want to talk about both the positive and negative sides of AI — specifically through three opportunities (Access, Insight, and Empowerment) and three threats (Bias, Transparency, and Privacy). I describe myself as a thoughtful optimist when it comes to AI adoption, so let's start with the opportunities.

Equitable access and the democratisation of knowledge

Remote diagnostics and telemedicine: access to expertise

First, AI can democratise access to expertise. Take the example of Millie, the Northern Territory breast-screening bus. It travels to remote areas, providing mammography to over 1,000 women across 20+ Aboriginal communities. Imagine a future where these women have access to the world's best diagnostic tools — right there in their communities, at no cost. That kind of access should not depend on whether you live in a city or a remote area, or whether you can afford a specialist.

Yes, developing and training AI models is expensive. But once trained and deployed at scale, the cost of querying these systems is negligible. So it becomes an affordable, scalable solution that can truly bridge equity gaps.

Translation and customisation of health information

During COVID-19, AI helped translate public health information into different languages for diverse communities. In a multicultural, multilingual country like Australia, this was critical. But beyond translation, AI can customise information to match a person's context, culture, and understanding — which is crucial for informed decision-making and, by extension, for democracy.

Enhanced decision-making and accelerated discovery

A few years ago, MIT showed that AI could detect breast cancer from mammography images five years earlier than conventional

methods — by identifying patterns too subtle for the human eye. And it's not just medical imaging. AI now processes molecular data at a scale impossible for humans.

In my own work, we've focused on liquid biopsies — measuring thousands of molecules in just a few drops of blood to detect or monitor cancer. These methods, powered by AI, allow us to track treatment response with a blood test, instead of costly, invasive procedures.

Whether it's genomics, microbiome data, wearables, or electronic health records, we're surrounded by diverse health data. When AI integrates these sources, it gives us a comprehensive view of health and disease — the vision behind precision medicine. We're not fully there yet, and some are sceptical. But I believe that without AI, this kind of personalised, holistic care simply isn't possible.

Personalised medicine, by nature, supports democratic principles. It ensures that people aren't disadvantaged just because their genome or biology deviates from the average.

Citizen empowerment and enhanced autonomy

AI also enables citizen empowerment — giving individuals tools and resources to take control of their health decisions. That includes personalised health management tools, decision support systems, and better access to understandable, relevant information.

These systems promote autonomy and support community building by helping people connect and share experiences. They also offer new ways to inform policymakers — giving communities a stronger voice.

But there are serious threats

Surveillance and privacy erosion

Of course, the flip side of AI-enabled access is surveillance. AI systems often depend on massive amounts of personal data. How that data is stored, sold, or used to monitor people poses a serious threat to democratic freedoms. The chilling effect — where people censor themselves because they feel watched — directly undermines free speech and open dialogue.

Bias in data and inequity in decisions

Bias is one of AI's most dangerous challenges. It comes from the data, and it gets baked into the decisions AI makes.

Here's a real example: I asked a generative AI model (GPT) to draw an image of a scientist. It produced a white man in a lab coat (A). I then asked for a university professor: again, a white man with a beard and an open-neck shirt (B). A research centre director? Same (C). A CEO? White man in a suit (D). I then asked it to draw an image of *me*, based on publicly available information. It showed a white woman in a suit, speaking outdoors (E). I asked, "How do you know I'm a woman?" The model replied: "Because you won an award in Women in AI APAC." Fair enough.

But then I asked it again: "Draw an image of me: a university professor, director of a centre, and CEO." The result? Back to an AI-generated image of a bearded white man in a suit (F), despite prior context indicating a female identity, reflecting the model's bias at the time (November 2024).

Clearly, the bias goes deeper than job titles or gender. And this isn't limited to

Figure 1: AI-generated illustrations created using GPT-4o based on a series of prompts.

illustrations. In our breast cancer blood test project, we trained a model using data from Eastern Europe. It performed well — until we tested it in Australia. It failed for women who weren't Caucasian. The model had learned bias from its training data, and couldn't generalise to Australia's diverse population. That's a critical equity failure — and it's happening in real-world applications.

Transparency and interpretability

If I ask you, "How does AI work?" and you say, "I don't know," you're not alone — even experts often don't know. That's because AI, especially deep learning, operates as a "black box:" millions or billions of values interacting in complex, nonlinear ways to generate an output.

Figure 2: Assoc Prof Fatemeh Vafaee

Yes, we have metrics to improve transparency in medicine, but accountability must go further. We need explainable systems that people — not just engineers — can understand and trust.

From fragile to antifragile

Let me end with a big-picture point: Right now, we're in a fragile state. We don't fully understand how AI systems behave, or how they will evolve. That unpredictability makes us vulnerable.

So how do we become antifragile — able to adapt, improve, and use AI responsibly?

We need measurement tools, transparency, auditing, and incident reporting. Only then can we truly evaluate AI's impact. And let me be clear: there is no path forward that doesn't include AI. We must embrace it — but we must do so responsibly.

SC: That was a fantastic talk about AI. I especially loved the example about bias — and how you still ended up as a white man in the end!

Q&A

Q1: That was a great discussion. My name is Essen, from the Office of the Chief Scientist and Engineer. I really appreciated hearing the different perspectives on how technology affects democracy — particularly around misinformation, equity, and bias.

We heard from Darren about how genomics and medicine are subject to bias, based on the data they're trained on — mainly from Caucasian populations. The same is true for AI. Recent research also shows that even language can introduce inequality. For example, English-speaking users receive significantly better responses from ChatGPT than people using minority languages.

I'd like to hear the Panel's thoughts on how this bias — whether in medicine, genomics, or AI — will impact democracy not just in Australia, but globally. Especially considering Australia's diversity, how do you see this playing out?

FV: Wearing my technical hat, I'd say there are real opportunities to address these issues using both regulation and technology. Many of the harms AI can cause — bias being a key example — can be mitigated when we combine thoughtful regulation with innovation.

In the case of bias, particularly related to ethnicity, we can begin by measuring the extent and nature of that bias in AI models. That helps users, developers, and regulators understand where corrections are needed. We're actively working on reducing data bias — both within the models themselves and in how we collect data across diverse populations.

But the solution isn't just about data collection. It's also about understanding how to reduce structural bias within the algorithms. Transparency and deeper technical insight are essential for making meaningful progress. This is a complex issue, but one we can address if we treat it as both a social and technical challenge.

Q2: This question is for all of you, especially Ed and Darren.

Given the rise of misinformation and declining trust in authority, do you think we need to explicitly teach critical thinking in schools? I have two teenage daughters, and while critical thinking is loosely part of the curriculum, it's not necessarily taught as a direct skill — like how to evaluate sources or understand what "truth" means online. Should this now be a core part of education in primary and secondary schools?

ES: The short answer is yes. It's essential for all the reasons you've mentioned.

But I do have a concern about where we tend to direct our solutions. Often, we put the burden on the victims of bad systems — like people being misinformed — to protect themselves. It's like saying, "We've

created a dangerous environment, now it's your job to wear armour."

I'll give you a real example. I once gave a big talk on facial recognition and how it risks creating a mass surveillance society. Afterward, a man raised his hand and said, "You've talked about legal and policy solutions, but I've got a better one." I got my pen ready, curious. He said, "I wear a beak in public." A beak? "Yes," he said. "It's modelled after an exotic South American bird, and it defeats all the facial recognition systems. If everyone wore beaks, we'd be fine."

Now, technically, he's not wrong — but I don't like the idea that people have to contort themselves, literally or metaphorically, to resist the harms of bad technology. Yes, teach critical thinking — wear a beak if you must — but more importantly, let's regulate and design technology to be human-centred from the start.

SC: You've got to give him full marks for lateral thinking.

DS: Just to add to that: while I'd love to think that teaching critical thinking will solve the problem, I'm actually not sure it will. From my experience communicating complex science to the public — and there's good research backing this — just giving people more or better information doesn't necessarily help. In fact, it can make things worse.

A lot of misinformation uptake is driven by belief systems and emotional reactions, especially during times of stress. Sometimes, when people are given facts that challenge their worldview, they double down instead of changing their minds.

So yes, teach critical thinking, but we also need to address the emotional and psychological roots of misinformation. It's not just a cognitive issue — it's a deeply human one.