View online

Centre for Media Transition

Hi there

Vulnerable but not silly

Welcome to our newsletter. This week Tamara tackles the tension between freedom of expression and tackling the spread of misinformation in Brazil where the country's Supreme Court has upheld a ban on X (formerly Twitter). Sacha delves into the controversial question around banning kids from social media, and Meta's unashamed admission that it's mining our data for its own AI machine and we can't stop them. Michael assesses the government's latest Combatting Misinformation and Disinformation Bill, tabled this week in Parliament. And I'm looking at research showing that maybe politically motivated, AI-generated 'deepfakes' aren't hitting the mark.

ACT Senator David Pocock was on the tools this last week, creating his own AI generated deepfakes – one of Prime Minister Anthony Albanese and the other of Opposition leader Peter Dutton, in a rare, albeit fake moment of bipartisanship, proposing a full ban on gambling ads. The Senator was attempting to make the point that the government needs to ban the use of AI-generated material ahead of the next federal poll to avoid our democracy being harmed.

Putting aside the political theatrics, it's worth thinking more about whether deepfakes have in fact impacted elections held in this year of elections around the world. Researchers at Oxford University and the University of Zurich conclude that the answer is actually, 'not so

much': 'early alarmist claims about AI and elections appear to have been blown out of proportion.'

The researchers noted that generative AI was anticipated to be cataclysmic – making it easier 'to create realistic but false or misleading content at scale, with potentially catastrophic outcomes for people's beliefs and behaviors, the public arena of information and democracy'. But when they examined if there had been an increased quantity of misinformation, an increase in the quality of misinformation and increased personalisation of misinformation, in nations where elections have been held – Pakistan, Bangladesh, India, Indonesia, Taiwan, Mexico, South Africa, the UK, Panama, and South Korea among many more – the instances and impact were significantly lower than anticipated. You can look here at a compendium of known incidents of AI-generated election-related misinformation in the nations named above.

Of course, AI is being used to generate misinformation in electoral processes. However, at this point – even in the US where the election campaign is well underway – the research shows that the level of misinformation is no greater than it usually is and that where AI-generated interference is noted, 'these efforts have not been fruitful'. This appears to be supported by the Alan Turing Institute in the UK, which looked at 112 elections held since 2023. It found the current impact of AI on specific election *results* is limited, but the threats show signs of damaging the broader democratic system. 'As of May 2024, evidence demonstrates no clear signs of significant changes in election results compared to the expected performance of political candidates from polling data.'

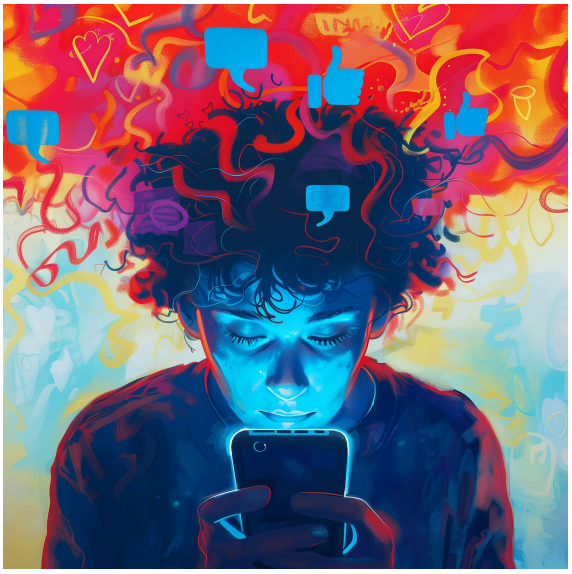
All of this led the Oxford University researchers to ask why the speculation about the damage from deep fakes and other AI-generated electoral content was so off the mark? They concluded that it's down to the fact humans are stubborn and not silly. New information (in the form of disinformation) might get them thinking, but it rarely translates to behavioural change. And with the overload of information that voters are presented with, AI-generated content isn't cutting through. The researchers also noted that 'voters seem to not only recognise excessively tailored messages but actively dislike them.'

They concluded that there are many more things around elections we ought to be deeply concerned about – politicians peddling lies, voter disenfranchisement and, last but not least, attacks on journalists.

And in completely unrelated but nonetheless intriguing news – Nine Chief Executive Mike Sneesby is stepping down and a 'global search' is underway for a new boss. It's been a turbulent time at Nine, with the departure of chair Peter Costello and the exit of some 100 journalists from the newspaper arm of the organisation.

Monica Attard
CMT Co-Director

Big tech, bad ethics



This week saw the announcement of major media reforms intended to hold big tech to account. It's been a long time coming, with privacy law, misinformation, online safety and the news media bargaining code among the many legal issues under ongoing review. Still, there were surprises. Expected reforms were absent; unexpected reforms made an appearance.

Top of the agenda is a social media ban for kids. And why not? There is a problem, clearly. Some have questioned how we verify the age of users, but isn't the answer obvious? TikTok, YouTube and Insta collect rather a lot of data. I'm confident they'd have a pretty good idea of their users' ages. The business model of digital platforms isn't complicated. One, collect data. Two, attract advertisers.

On that note, this week the government revealed how its privacy law reform is progressing. Yesterday, a bill was introduced to Parliament containing the first tranche of reforms, including the introduction of a statutory tort for serious invasions of privacy. This is a major step forward, but as Privacy Commissioner Carly Kind says, there is a lot more to be done: 'Further reform of the Privacy Act is urgent.'

Indeed. Also, this week Meta made unedifying admissions about the data it's been scraping to train its AI, with no opt-out available for users. Here's a question for Meta: apart from the data of users, are you also scraping the data of non-users of your services to train your AI? In other words, take someone who has never had an account with Facebook, Instagram, WhatsApp, Threads, Messenger, or any other Meta offering. Nonetheless, that person appears in users' photos and posts. Is Meta scraping that non-user's data? If the answer is yes, it's time for serious self-reflection. And for tough law. Let's see what the second tranche of privacy reform looks like.

For all these issues, a fundamental shift is needed. Let's put the responsibility on digital platforms and switch from a caveat emptor (buyer beware) approach to a caveat venditor (seller beware) approach. I've previously argued this in relation to privacy and, in a new paper, in relation to algorithms and AI. On this line of thinking, it's the responsibility of digital platforms to ascertain the age of users, and to ensure data collection and use is ethical. Further, I argue that the law should encourage 'light patterns', instead of dark patterns that manipulate users into acting against their own best interests.

Sure, let's encourage innovation, but not at the expense of our autonomy and democracy.

Sacha Molitorisz
Senior Lecturer - UTS Law

At your discretion

The release of the exposure draft of the Combatting Misinformation and Disinformation bill in July last year was met with widespread criticism – less a chorus than a cacophony. Chief amongst many fears was the potential threat to free speech from new powers the bill would grant to the Australian Communications and Media Authority. More than a year after consultation on the exposure draft closed, the government has finally introduced a revised bill to parliament.

The revised bill includes several changes that should allay some of the concerns over free speech. Most significantly, clause 67 proscribes the imposition of any rule that would require a platform to remove content or user accounts, except where such a rule would address 'inauthentic behaviour' – the use of automated systems or coordinated action to deceive users (noting that this is quite narrowly defined). Conversely, nothing in the bill prevents a platform from removing content or user accounts. Content-moderation policies are therefore left to the discretion of digital platforms. This change gives substance to the government's claims that the legislation is designed to promote transparency and accountability over platform systems and processes.

The revised bill also includes minimum requirements that apply outside any code registered or standard imposed by ACMA. These include obligations to conduct risk assessments and publish the results, to publish a policy or provide information on the platform's approach to misinformation, and to publish a media literacy plan explaining how the platform empowers users to identify and respond to misinformation. ACMA may augment these requirements by imposing rules relating to these areas as well as to user complaints and record keeping. These new provisions represent a significant improvement in the ability of the legislation to promote platform accountability and transparency.

Unfortunately, other problems remain unresolved. As the exposure draft did before it, the revised bill undermines its ability to promote accountability by setting too narrow a scope, both through restrictive definitions of misinformation and disinformation that require reasonable likelihood of serious harm and through the exclusion of professional news and private messaging. A narrow scope protects freedom of expression from government overreach. But it also limits accountability by excluding the majority of platform content moderation decision-making from its scope. If a platform chooses to remove my post (or say, a professional news story), but that post does not surpass the threshold of serious harm, then this bill will not ensure I have recourse to dispute the platform's decision. It is outside ACMA's powers to impose a rule requiring a platform to provide me with such recourse.

Leaving content moderation to the discretion of platforms means that platforms are effectively the arbiters of truth. Better than government, you might say, and I would agree. But platforms undeniably have power over our freedom of expression, and there is no guarantee that they will act to protect and promote it. But this is as much an objective of the legislation as mitigating harm.

As I've noted before, this might seem like an unfortunate but inexorable dilemma. But it isn't. We can maximise platform accountability and minimise the threat of overreach. Specifying that the discretion to moderate content remains with platforms means there is no need to be unduly restrictive over the scope of content or services to which the bill applies.

Michael Davis
CMT Research Fellow

Come to Brazil! Except you, X

Last month, Brazil's Supreme Court ordered a nationwide block of X for its failure to appoint a local legal representative. While it's reassuring to know that even billionaire-funded social media giants aren't above the law, the court's decision highlights that there is a fine line between platform regulation efforts and censorship.

The decision was somewhat overshadowed by the months-long public feud between X owner Elon Musk and

Brazilian Supreme Court Justice Alexandre de Moraes, who handed down the block.

Moraes had included X and several other platforms in investigations targeting 'digital militias'. These were accounts accused of spreading disinformation and hate speech in the wake of right-wing President Jair Bolsonaro's election defeat and the subsequent storming of government buildings in January 2023.

Nonetheless, in April of this year, Musk publicly defied Moraes's orders to remove some of these accounts. He referred to the Supreme Court Justice as a 'dictator', called for his impeachment and joked that he looked like the spawn of Voldemort and a Sith Lord. Moraes, meanwhile, accused Musk of obstructing Brazilian justice and disrespecting its sovereignty.

But this isn't just a spat between a CEO and a judge. Brazil, like many other countries, has been grappling for years with how to balance the principles of free expression and media pluralism with rules that address illegal and harmful content.

Brazil is also home to millions of dedicated internet users, and over 20 million of them were active on X. No doubt many of us have encountered the phrase 'Come to Brazil!' somewhere in our internet travels, a now famous internet meme born from the consistency with which Brazilian fans wrote it under celebrity posts. Kaitlyn Tiffany has written a memorable piece in *The Atlantic* about how these users have played a strong role in developing online fan cultures, with some now feeling resentful towards both Musk and the Brazilian judiciary for the loss of their accounts.

Although a portion of those X users have flocked to BlueSky, some critics view the court's decision as an attack on free speech and political opposition; thousands of Bolsonaro supporters recently took part in a pro-free speech protest in São Paolo.
Not only that, but people caught accessing X in Brazil – even through a VPN – face a daily fine of around 50,000 BRL (almost US$9,000).

Supporters of the ban say it was an 'inevitable outcome' of the platform's continued non-compliance with national regulations, and that this may add to the 'growing mood' that social media companies are not above the law.

Either way, the case is a complex one that highlights the difficulties in balancing competing interests, and many will be watching with great interest to see how it pans out.

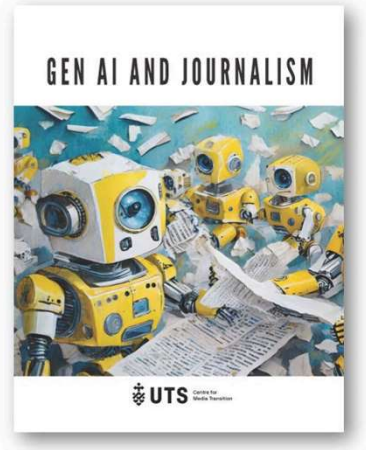
Tamara Markus
CMT Researcher

We hope you have enjoyed reading this edition of the *Centre for Media Transition newsletter* | *Social media spotlight, fakes and Brazil boots out X* | *Issue 17/2024*
ISSN 2981-989X

This serial can be accessed online here and through the National Library of Australia.
Please feel free to share our fortnightly newsletter with colleagues and friends!
And if this was forwarded to you, please subscribe by clicking the button below:

Subscribe

Please visit our website for more information about the Centre .



The Centre for Media Transition and UTS acknowledges the Gadigal and Guring-gai people of the Eora Nation upon whose ancestral lands our university now stands.
We pay respect to the Elders both past and present, acknowledging them as the traditional custodians of knowledge for these places.