

Political deepfakes and the new laws in NSW

Lenny Roth and Tom Gotsis

1. Introduction

There are international concerns about the potential impacts of artificial intelligence (AI) on democracy.¹ Among these concerns is the proliferation of AI-generated deepfakes that are designed to mislead voters around elections. Deepfakes have ‘been deployed to influence elections in India, Indonesia, Mexico, Pakistan, Slovakia, Ukraine, the USA and Taiwan.’² Some countries have introduced legislative reforms to combat political deepfakes. In March 2026 NSW became the second Australian jurisdiction to enact legislative reforms. This paper examines the potential adverse impacts of political deepfakes, discusses law reforms in Australia and overseas, and outlines other measures that can be used to combat political deepfakes.

2. What are political deepfakes?

A 2024 systematic literature review found that there is as yet no universally accepted definition of the term ‘deepfake’.³ This paper defines a ‘deepfake’ as digital content which has been created with artificial intelligence (AI) that depicts a person doing or saying something they did not do or say.⁴ It defines a ‘political deepfake’ as a deepfake that depicts: (a) a politician, or (b) a person expressing a view about a politician or political issue. Political deepfakes most typically take the form of videos, images and audio, or a combination of those forms. As shown in Table 1, political deepfakes can be realistic or obviously fake, and their sentiment can be either positive or negative.

Table 1: Content-based classification of political deepfakes⁵

Type	Features
Realistic and negative	Pose a direct threat to a politician’s wellbeing and reputation For example, they may show a politician being sexually explicit or endorsing hate speech or violence They can sway voters through deliberate deception because they are extremely realistic
Realistic and positive	Aim to glorify a politician through fabricated positive moments For example, they may show a politician deliver a powerful speech they never gave, do something virtuous, or battle adversity They have the potential to sway voters through deliberate deception because they are extremely realistic
Unrealistic and negative	Depict a politician negatively using obviously fictional scenarios (or are marked as being AI-generated) For example, a politician may be depicted as acting ineptly or foolishly (such as dancing in support of nuclear power). Often described as a form of political satire Potential to sway voters does not depend on extreme realism or deliberate deception, but on satire and negative associations

¹ Christopher Summerfield et al, ‘[The impact of advanced AI systems on democracy](#)’, *Nature Human Behaviour*, pp. 2420–2430. For discussion in Australia, see Senate Select Committee on Adopting Artificial Intelligence, [Select Committee on Adopting Artificial Intelligence – Interim Report](#), Parliament of Australia, October 2024, Ch 2.

² C. Summerfield, *The impact of advanced AI systems on democracy*, p 2423-2424.

³ A Birrer and N Just, ‘[What we know and don’t know about deepfakes: An investigation into the state of the research and regulatory landscape](#)’, *New Media and Society*, 27(12) 2024.

⁴ See also: T Gotsis, [Sexually explicit deepfakes and the criminal law](#), NSW Parliamentary Research Service, 2025, pp. 5-6. The use of AI means that deepfakes are more technologically sophisticated and convincing than ‘cheapfakes’, which can be made using simple software functions, such as copy and paste.

⁵ Based on: M Wack et al, ‘Scrutinizing the many faces of political deepfakes’, *Tech Policy Press*, 18 November 2025. Accessed at: <https://www.techpolicy.press/scrutinizing-the-many-faces-of-political-deepfakes/>

Type	Features
Unrealistic and positive	<p>Depict a politician positively in an obviously fictional scenario (or are clearly marked as being AI-generated)</p> <p>For example, a politician may be shown as a superhero or revered historical figure</p> <p>Potential to sway voters does not depend on extreme realism or deliberate deception but on developing a relationship between voters and a fake heroic personality</p>

There is no single authoritative source of data on the prevalence of deepfakes, with some academics noting that any existing data should be viewed cautiously, as this is an emerging area of research.⁶ Nevertheless, both research institutions and private companies are attempting to identify the number and type of deepfakes that are online and track future trends. In 2023 a university and private sector collaboration started collecting data for the Political Deepfakes Incidents Database.⁷ As at 24 August 2026, the database included records of 2,436 political deepfakes.⁸

One Australian political deepfake was located in the database. It was a sexualised image of Georgie Purcell, a member of the Victorian Parliament, which was published by a media outlet.⁹ There are other Australian political deepfakes that have been reported in the media which do not appear in the database. For example, a deepfake of former Queensland Premier Annastacia Palaszczuk in which she described Queensland as having ‘massive debt’ and ‘huge unemployment’; and a deepfake of then federal opposition leader Peter Dutton dancing in support of nuclear power plants¹⁰

3. Potential adverse impacts

The potential adverse impacts of political deepfakes relate to their effects on elections, politicians and political communication.

3.1 Elections

Some political deepfakes seek to influence people’s voting preferences through the use of deception.¹¹ Such deepfakes have the potential to create uncertainty and doubt about all political information that is presented to voters before an election.¹² There can also be a national security dimension to this issue, where political deepfakes have been created by a foreign state to interfere in

⁶ A Birrer and N Just, *What we know and don’t know about deepfakes*, 2024.

⁷ A Azzo, [Tracking political deepfakes: New database aims to inform, inspire policy solutions](#), Centre for Advancing Safety of Machine Intelligence, 26 January 2024.

⁸ This figure excludes 212 instances where it was unclear whether material was a deepfake or a cheapfake, and 61 instances where the material was real (i.e., a false accusation of a deepfake). The database was searched by selecting: Filter by type/type/is/deepfake.

⁹ A Sarkar, ‘Photoshop owner rejects TV channel’s claim that AI to blame for doctoring female MP’s image to be more revealing’, *Irish Independent*, 30 January 2024. Accessed at: <https://www.independent.ie/world-news/photoshop-owner-rejects-tv-channels-claim-that-ai-to-blame-for-doctoring-female-mps-image-to-be-more-revealing/a/129063282.html>

¹⁰ S Richards, ‘As deepfakes advance with technology, there are concerns they could become a ‘threat to democracy’, *ABC News*, 8 October 2024. Accessed at: <https://www.abc.net.au/news/2024-10-08/deepfake-videos-queensland-election-tiktok-misinformation/104360096>. It was claimed that these deepfakes were satire because they were clearly marked as being deepfakes.

¹¹ M Wack et al, *Scrutinizing the many faces of political deepfakes*. For a discussion of notable instances where deepfakes have been deployed during elections, see: E Cervini and M Carro, [An overview of the impact of GenAI and Deepfakes on global electoral processes](#), *ISPI Analysis*, March 2024.

¹² A Birrer and N Just, *What we know and don’t know about deepfakes*, pp. 6,819-6838.

the elections of other states.¹³ In these ways, political deepfakes threaten to undermine elections and democracy by limiting the ability of people to exercise their free and informed choice as voters and by reducing their confidence in the integrity of elections and democratic institutions.¹⁴

The threat to elections and democracy can be seen in the events surrounding the 2025 Irish presidential election, which was the subject of 172 instances of political misinformation and disinformation attacks published on social media in the 6 weeks leading up to polling day.¹⁵ The attacks included deepfake news reports which falsely showed candidate Catherine Connolly withdrawing from the election. The deepfake news reports also falsely stated that the election had been cancelled and another candidate had won by default.¹⁶ Ms Connolly reported the deepfakes to the relevant social media platforms and lodged a formal complaint with the Electoral Commission, after which the deepfakes were removed from some platforms.¹⁷ She also issued a media statement denouncing the deepfake as a 'disgraceful attempt to mislead voters and undermine our democracy' and reassuring the electorate of her candidacy.¹⁸ Ms Connolly won the presidential election.

3.1.1 Evidence on the impact of political deepfakes on election results

Stakeholders have pointed out that the 'current research on the impact of AI on elections ... has not uncovered any clear impact of AI-generated deepfakes on actual voting outcomes'¹⁹ or election results.²⁰ Those stakeholders include the Australian Electoral Commission, which has said: 'The impact is hard to quantify but there hasn't been any evidence to date that the use of AI in election communication has been the determining factor in election results.'²¹ It has been suggested, however, that although 'exactly how many voters could be misled by a deepfake remains unclear ... if marginal seats were targeted during an election, even swaying as few as 100 voters could be impactful.'²² It has also been suggested that the lack of clear evidence is 'no cause for complacency'.²³

¹³ See: C Leis, 'Addressing the role of deepfake technology in foreign interference in US elections', University of Virginia National Security Data and Policy Institute, 2025. Accessed at: https://nspcbatten.org/wp-content/uploads/sites/2/2025/07/C.-Leis_Research-Placemat.pdf

¹⁴ A Ray, *Disinformation, deepfakes and democracies: The need for legislative reform*, *UNSW Law Journal*, 2021, 44(2), p. 987. See also: M Platow et al, *Information resilience: How misinformation undermines social cohesion, trust and democracy*, Australian Government Office of the Chief Scientist, 2026, Ch 1.4.

¹⁵ Institute for Strategic Dialogue, 'Irish Presidential Election 2025: Renewed attacks on election integrity and repeated platform failures', 31 October 2025. Accessed at: <https://www.isdglobal.org/digital-dispatch/irish-presidential-election-2025-renewed-attacks-on-election-integrity-and-repeated-platform-failures/>

¹⁶ J Liggett, 'Disgraceful' deep-fake AI video condemned by presidential candidate', *BBC*, 23 October 2025. Accessed at: <https://www.bbc.com/news/articles/czxkn504lqpo>

¹⁷ T McNally, 'Malicious deepfake': Condemnation of AI video falsely showing Catherine Connolly ending campaign', *Irish Examiner*, 22 October 2025. Accessed at: www.irishexaminer.com/news/politics/arid-41728776.html

¹⁸ Connolly for President, 'Connolly condemns 'malicious deep-fake' video', 22 October 2025. Accessed at: <https://www.catherineconnollyforpresident.ie/press-release-22-10-25-deepfake>

¹⁹ O Carpenter-Zehe, 'Irish election deepfake AI video shines light on lack of EU-wide rules', *euobserver*, 27 October 2025. Accessed at: <https://euobserver.com/3226/irish-election-deepfake-ai-video-shines-light-on-lack-of-eu-wide-rules/>

²⁰ S Stockwell, *AI-Enabled influence operations: Threat analysis of the 2024 UK and European elections*, The Alan Turing Institute and Centre for Emerging Technology and Security, September 2024, p. 27. See also: M Labuz and C Nehring, *On the way to deepfake democracy? Deepfakes in election campaigns in 2023?*, *European Political Science*, 23, 26 April 2024, pp. 454–473. Labuz and Nehring noted that whether political deepfakes have 'a decisive impact on voting behaviour... is extremely difficult, if possible at all, to be measured'.

²¹ Australian Electoral Commission, 'AI & Elections', 25 September 2025. Accessed at: <https://www.aec.gov.au/Elections/electoral-advertising/ai-and-elections.htm>

²² A Ray, *Disinformation, deepfakes and democracies*, p. 988.

²³ S Stockwell et al, *AI-enabled Influence Operations: Safeguarding Future Elections*, The Alan Turing Institute and Centre for Emerging Technology and Security, November 2024, p. 66.

3.1.2 Loss of public confidence in elections

There is some evidence, and broadly held concerns, that political deepfakes have the potential to erode public confidence in elections. A threat analysis of elections held in the United Kingdom and Europe in 2024 found that the campaign environment was being polluted by fake information that was generated by political deepfakes and social bots. This led to confusion over whether information about the election was authentic or fake, which undermined trust in the information environment surrounding elections.²⁴ The Australian Senate's Select Committee on Adopting Artificial Intelligence also noted the risk that AI generated political disinformation '... can lead to an erosion of public trust in and engagement with politics more generally'.²⁵

3.2 Effects on politicians

3.2.1 Reputational harm

Concerns have been raised for some time about the scope for political deepfakes to cause reputational harm to politicians.²⁶ It has even been noted that, as well as potentially damaging their electoral prospects, in extreme cases the reputational harm caused to politicians could be so significant that it leads to abuse, threats or violence.²⁷

3.2.2 Distress and psychological harm

Victims of deepfakes can, depending on the nature of the deepfake, suffer from distress and psychological harm. A February 2026 article suggests that deepfake victimisation is a new category of digital trauma that warrants clinical attention and policy actions.²⁸ The article notes that acute anxiety, insomnia, intrusive recollections, hypervigilance and social withdrawal are frequently experienced by victims and, in extreme cases, the impacts of deepfake victimisation can result in clinically significant acute stress disorder and post-traumatic stress disorder.²⁹

A particularly harmful form of deepfake victimisation involves sexually explicit deepfakes being used to intimidate, demean and blackmail politicians, particularly (although not exclusively) female politicians.³⁰ The literature indicates that victims of sexually explicit deepfakes suffer from emotional, psychological, personal, social and professional distress.³¹ Several female politicians in the United Kingdom have already been targeted in this way, with one victim noting that 'none of this is about

²⁴ S Stockwell, *Threat analysis of the 2024 UK and European elections*, p. 28.

²⁵ Australian Parliament, [Select Committee on Adopting Artificial Intelligence \(AI\): Interim Report](#), October 2024, p. 9.

²⁶ See, for instance, B Chesney and Danielle Citron, [Deepfakes: A looming challenge for privacy, democracy and national security](#), *Californian Law Review*, 107, 2019, p. 1774.

²⁷ S Stockwell, *Threat analysis of the 2024 UK and European elections*, pp. 8-9 and L Kharvi, [Understanding the impact of AI-generated deepfakes on public opinion, political discourse, and personal security in social media](#), *IEEE Security and Privacy Magazine*, July 2024, p 4.

²⁸ M Umar, [Digital trauma: deepfake victimization and AI-generated violence](#), *The Lancet: Psychiatry*, 13(2), February 2026, pp. 89-90.

²⁹ M Umar, *Digital trauma*, pp. 89-90.

³⁰ In NSW, the creation or distribution of a sexually explicit deepfake is a criminal offence under sections 91PA or 91Q of the *Crimes Act 1900*, whereas as the publication of a sexualised image (which is not explicit) is not a criminal offence.

³¹ A Flynn, J Clough and T Cooke, 'Disrupting and preventing deepfake abuse: Exploring criminal law responses to AI-facilitated abuse', Chapter 29, p 587-589 in A Powell et al (eds), *The Palgrave Handbook of Gendered Violence and Technology*, Great Britain, 2021. See also: T Gotsis, [Sexually explicit deepfakes and the criminal law](#), NSW Parliamentary Research Service, 2025, pp. 8-10.

sexual pleasure, it's all about power and control'.³² In Malaysia 10 politicians, including at least 4 male politicians, were blackmailed with sexually explicit deepfakes, with the offenders demanding \$US100,000 in cryptocurrency.³³ A 2025 article found that 26 members of the U.S. Congress were depicted in more than 35,000 instances of deepfake content on pornographic websites, and 25 of these members were women.³⁴

3.2.3 Chilling effect on aspiring and existing female politicians

Concerns have been raised that political deepfake victimisation may have a chilling effect on the democratic participation of women by deterring them from becoming politicians, remaining in politics, advocating for policy reform, or seeking leadership positions. The Inter-Parliamentary Union views deepfake victimisation of female politicians as falling into:

... a wider trend of abuse targeting women in the public scene. Sexism, harassment, and violence against women MPs are quickly becoming a global problem impeding gender equality and undermining the foundations of democracy.³⁵

3.3 Broader effects on political communication

In western democracies, political communication involves an effort to influence the exercise of people's free and informed choice as electors. In Australia, with the exception of the ACT and South Australia, there is no legal requirement that the effort to influence voters during elections must be based on truth.³⁶ It has been observed, however, that political communication is 'most functional when debates build from a foundation of shared facts and truths supported by empirical evidence.'³⁷

Australia's political history has frequently shown that building a foundation of shared facts and truths can be challenging.³⁸ Those challenges are heightened by the combined effect of social media and cognitive biases that cause people to believe and share information that accords with their pre-existing beliefs and to ignore information which does not.³⁹ Academics have suggested that political deepfakes have the potential to further exacerbate those challenges, which will increase political polarisation,⁴⁰ confuse people as to what constitutes a fact, and:

³² See, for instance: J Waterson, 'British female politicians targeted by fake pornography', *The Guardian*, 2 July 2024. Accessed at: <https://www.theguardian.com/technology/article/2024/jul/01/british-female-politicians-targeted-by-fake-pornography>

³³ 'Malaysian politicians targeted in AI deepfake blackmail campaign', *Folio3*, September 2025. Accessed at: <https://www.folio3.ai/ai-pulse/malaysian-politicians-targeted-in-ai-deepfake-blackmail-campaign>. See also: M Mustafa, Malaysia grapples with AI legal grey zone as deepfake porn blackmail targets lawmakers, *The Straits Times*, 23 September 2025. Accessed at: <https://www.straitstimes.com/asia/se-asia/malaysia-grapples-with-ai-legal-grey-zone-as-deepfake-porn-blackmail-targets-lawmakers>

³⁴ M Clarice, 'Deepfakes, AI, and the new war against women in politics', *WePro*, 8 January 2025. Accessed at: <https://wepro.digital/article/deepfake>

³⁵ Inter-Parliamentary Union, 'Dangers of deepfakes for parliamentarians', 19 February 2024. Accessed at: <https://www.ipu.org/news/news-in-brief/2024-02/dangers-deepfakes-parliamentarians>

³⁶ A 2026 report found that '81% of Australians support a legislated requirement that political advertisements meet standards of truthfulness, and only 5% oppose': Democracy Counts, *The State of Australia's Democracy*, 2026, p 3.

³⁷ B Chesney and D Citron, *Deepfakes: A looming challenge*, p. 1,777.

³⁸ There are many examples from Australia's recent political history which can be used to illustrate the inherent difficulty in building a foundation of shared facts and truths. They include debates about national identity, reconciliation, immigration, same-sex marriage, free speech, protests, climate change, negative gearing, public school funding, wealth inequality, immunisation and pandemic lockdowns.

³⁹ B Chesney and D Citron, *Deepfakes: A looming challenge*, pp. 1764–1768 and 1786. See also: M Platow et al, *Information resilience*, 2026.

⁴⁰ A Amin et al, *The influence of social media deepfake images on political ideology and polarization: the mediating roles of cognitive load and confirmation bias*, *Journal of Visual Literacy*, 44, 2025, pp. 321-339.

... allow individuals to live in their own subjective realities, where beliefs can be supported by manufactured 'facts.' When basic empirical insights provoke heated contestation, democratic discourse has difficulty proceeding. In a marketplace of ideas flooded with deep-fake videos and audio, truthful facts will have difficulty emerging from the scrum.⁴¹

Political deepfakes could also 'contribute to a general climate of uncertainty and doubt', which would reduce overall trust in and engagement with political information.⁴² The climate of uncertainty and doubt created by deepfakes further damages political discourse by making it easier for politicians to deny the truth. Chesney and Citron have called this phenomenon 'The Liar's Dividend', which can operate in 2 distinct ways. First, a politician might create doubt about an accusation made against them by creating and distributing a deepfake that appears to contradict the claim. Second, as the public becomes more aware of the threats posed by political deepfakes, they become more sceptical of news. This heightened scepticism makes it easier for a politician to lie by falsely claiming that real evidence which casts them in a negative light is 'fake news'.⁴³

Not all deepfakes, however, are detrimental to political communication. Indeed, some deepfakes can make positive contributions to political communication. For instance, a politician could make 'videos of *themselves* speaking in different languages in efforts to appeal to a greater share of the voting base'.⁴⁴ Political deepfakes which are satirical, rather than deceptive, could arguably make a distinct and significant contribution to political discourse.⁴⁵ Satire, it has been said, is 'the most important form of public humour' because it is 'designed to make society examine itself critically and confront its deficiencies'.⁴⁶

4. Law reform in Australia and overseas

4.1 New laws in NSW

4.1.1 Electoral law reforms

In March 2026 the NSW Government introduced new offences relating to AI-generated electoral material as part of reforms to the *Electoral Act 2017* in the lead up to the 2027 state election.⁴⁷ The government explained that the purpose of the new offences is:

...to safeguard New South Wales elections from the spread of misinformation by artificial intelligence, which may be used to manipulate public opinion or unfairly target specific groups of voters, potentially undermining the fairness and integrity of the entire election process.⁴⁸

There are 2 new offences. First, section 189A makes it an offence for a person to, during an election period, publish or distribute digitally generated electoral material that contains a depiction of a

⁴¹ B Chesney and D Citron, *Deepfakes: A looming challenge*, p. 1778.

⁴² A Birrer and N Just, *What we know and don't know about deepfakes*, pp. 6,819-6838.

⁴³ B Chesney and D Citron, *Deepfakes: A looming challenge*, p. 1785. See also: JA Goldstein and A Lohn, 'Deepfakes, elections and shrinking the Liar's Dividend', Brennan Centre For Justice, 23 January 2024. Accessed at: <https://www.brennancenter.org/our-work/research-reports/deepfakes-elections-and-shrinking-liars-dividend>

⁴⁴ A Ray, *Disinformation, deepfakes and democracies*, p. 1,007.

⁴⁵ A Ray, *Disinformation, deepfakes and democracies*, p. 1,007.

⁴⁶ J Parker, *A licence for the satirist? The failure of Australian defamation law to protect satire in political media* UNSW LawJLStus, 2021, p. 21. Parker notes that satire 'has remained relevant across varying political systems and cultures in the Western World since its origin in Graeco-Roman civilisation'.

⁴⁷ *Electoral Legislation Amendment (Elections) Act 2026* (NSW).

⁴⁸ J. Aitchison, New South Wales, *Parliamentary Debates*, Legislative Assembly, 17 March 2026, p 38.

simulated person performing an act that the real person depicted did not perform. Second, section 189B makes it an offence for a person to, during an election period, publish or distribute digitally generated electoral material that depicts a digitally created person, event or place that a reasonable person would think is an actual person, event or place unless the material contains a statement that the material is digitally generated.

Several terms are defined in the Act. Digitally generated electoral material:

- (a) means electoral material containing audio visual, visual or audio content that is generated substantially, or modified or altered significantly, by artificial intelligence, but
- (b) does not include
 - (i) a cartoon or animated drawing, or
 - (ii) electoral material that is of a kind, or generated in a way, prescribed by the regulations.⁴⁹

The election period means the period from the date of the issue of the writ for the election until 6 pm on election day.⁵⁰ Electoral material means anything containing 'electoral matter'. 'Electoral matter' means any matter that is intended or capable of affecting the result of any election held or to be held or that is intended or capable of influencing an elector in relation to the casting of his or her vote.⁵¹ A simulated person includes a person depicted in digitally generated electoral material that 'purports to be a depiction of a particular real person'.⁵²

It is a defence to the first offence outlined above if the publication occurred with the written consent of the real person depicted as the simulated person. It is a defence to both the first and second offences if the defendant establishes that he or she (a) took no part in determining the content of the digitally generated electoral material, and (b) could not reasonably be expected to have known the material contravened the relevant provision. In the case of an individual, the maximum penalty for the first offence outlined above is a fine of \$6,600 and/or imprisonment for 6 months, and the maximum penalty for the second offence is a fine of \$2,200 and/or imprisonment for 6 months.

4.2.1 Constitutional issues

There was significant parliamentary debate about whether the NSW offences would breach the implied freedom of political communication in the Australian Constitution. The Opposition argued that the new offences would breach the implied freedom and moved amendments to create an additional defence of substantial truth or fair comment.⁵³ The government did not support the amendments, arguing that the laws were proportionate.⁵⁴

In a 2021 journal article on political deepfakes laws, Ray suggested that even false speech may be protected by the implied freedom, and he concluded that the key issue is ensuring that the laws 'are suitable, necessary and adequate in their balance'.⁵⁵ He proposed an offence of publishing an altered image if it 'contains a statement regarding electoral matters that is inaccurate or misleading to a

⁴⁹ *Electoral Act 2017* (NSW), s 189A(4). The government explained that it 'does not intend to prescribe certain types of excluded electoral material in a regulation at this time: J. Aitchison, New South Wales, *Parliamentary Debates*, Legislative Assembly, 25 March 2026, p 43.

⁵⁰ *Electoral Act 2017* (NSW), s 4.

⁵¹ *Electoral Act 2017* (NSW), s 4.

⁵² *Electoral Act 2017* (NSW), ss 189A(4), 189B(4).

⁵³ A. Henskens, New South Wales, *Parliamentary Debates*, Legislative Assembly, 25 March 2026, pp 1-3..

⁵⁴ J. Aitchison, New South Wales, *Parliamentary Debates*, Legislative Assembly, 25 March 2026, p 41.

⁵⁵ A. Ray, *Disinformation, deepfakes and democracies*, p. 1005-1007.

material extent.’⁵⁶ He also proposed defences including if it was published for the purpose of ‘education, comedy, or satire’ and it was ‘identified as an altered image’.

4.2.3 Other law reforms

In 2025 new offences were inserted into the *Crimes Act 1900* (NSW) to prohibit the creation and distribution of sexually explicit deepfakes.⁵⁷ These were intended to supplement offences enacted at the federal level in 2024 of using a carriage service to transmit sexually explicit deepfakes.⁵⁸

4.2 New laws in South Australia

South Australia is the only other Australian jurisdiction that has reformed its electoral laws to address political deepfakes. These reforms were enacted in the context of existing truth in political advertising laws. Following amendments in 2024 a new section 115B was added to the *Electoral Act 1985* to prohibit the distribution of an artificially generated electoral advertisement that contains a depiction of a simulated person performing an act that the real person depicted in the depiction did not perform.⁵⁹ An ‘electoral advertisement’ is an advertisement containing matter calculated to affect the result of an election. There are similar defences to those in NSW. In the case of an individual, the maximum penalty for the offence is a fine of up to \$5,000.

4.3 National law reform proposals

4.3.1 Electoral law reform proposals

In November 2024 the Australian Government introduced a bill to enact truth in political advertising laws, which included specific provisions on deepfakes.⁶⁰ The bill was not debated and lapsed on dissolution in March 2025. Under the bill, a person would be prohibited from authorising during an election period the communication of ‘authorisable electoral matter’ that contained a purported factual visual or audio depiction of an election candidate, which was inaccurate and misleading to a material extent.⁶¹ Authorisable electoral matter mainly included advertisements in which all or part of the distribution or production was paid for.⁶² Electoral matter would not include a communication for a dominant purpose that is a satirical, academic, educative or artistic purpose.⁶³ Breaches of the these provisions would attract a civil penalty of \$330,000.

4.3.2 Other law reform proposals

In November 2025 Senator David Pocock introduced a private members bill to regulate deepfakes in any context, not just elections. The bill proposes establishing a complaints and enforcement scheme for the non-consensual sharing of artificially generated audio or visual content that depicts a person’s face or voice.⁶⁴ The scheme includes civil penalties, and powers for the eSafety Commissioner to issue removal notices to social media service providers and hosting service providers. The bill was debated in Senate in June 2026 but, at the time of writing, has not been voted on.

⁵⁶ A. Ray, *Disinformation, deepfakes and democracies*, p. 1011.

⁵⁷ *Crimes Act 1900* (NSW), s 91PA introduced by *Crimes Amendment (Intimate Image and Audio Material) Act 2025*.

⁵⁸ See *Criminal Code Act 1995* (Cth), ss 474.17A, 474.17AA.

⁵⁹ *Electoral (Miscellaneous) Amendment Act 2024* (SA).

⁶⁰ *Electoral Legislation Amendment (Electoral Communications) Bill 2024* (Cth)

⁶¹ See proposed ss 321M, 321MA.

⁶² See proposed s 321JA(2).

⁶³ See note to proposed s 321JA(2) referring to s 4AA(5).

⁶⁴ *Online Safety and Other Legislation Amendment (My Face, My Rights) Bill 2025* (Cth). See also D. Pocock, *New Bill to protect identity in deepfake future*, 24 November 2025.

4.4 Law reforms in other countries

4.4.1 Electoral law reforms

In the United States, 31 states have enacted political deepfake laws.⁶⁵ Two states—Minnesota and Texas — prohibit the publication of political deepfakes a certain number of days prior to an election. Maryland prohibits the dissemination of political deepfakes with intent to influence a voter's decision (but is not limited to an election period). The other 28 states require disclosures on the media, stating if it contains a deepfake.⁶⁶ Some state political deepfake laws have been challenged in the courts on various grounds including that they violate the right to free speech in the US Constitution.⁶⁷

In Canada, in 2025 the Manitoba provincial government introduced an offence of, during an election period, knowingly distributing a deepfake with the intention of affecting the results of an election or undermining public confidence in the results or administration of an election.⁶⁸ The Commissioner of Elections can issue a stop notice to a person committing this offence and failure to comply can lead to a fine of up to \$20,000 for each day of non-compliance. In June 2026 the federal government created a new offence of creating or distributing a deepfake of an election candidate with intent to mislead.⁶⁹ There is an exception for parody or satire.

Laws banning political deepfakes have also been enacted in South Korea⁷⁰ and Singapore.⁷¹ The Singapore laws empower the Returning Officer to issue a written direction to require a social media service or internet access service to remove electoral advertising that breaches the provision, and failure to comply is an offence with a maximum penalty of \$1 million.⁷²

4.4.2 Other law reforms

In 2022 the European Union (EU) enacted the *Digital Services Act*, which requires providers of 'very large online platforms' and 'very large online search engines' to mitigate systemic risks stemming from their services including any negative effects on electoral processes.⁷³ Such mitigation measures may include ensuring that a generated or manipulated image, audio or video that appreciably

⁶⁵ National Conference of State Legislatures, 'Deepfakes in elections and campaigns'. Accessed at: <https://www.ncsl.org/elections-and-campaigns/artificial-intelligence-ai-in-elections-and-campaigns>. For proposed federal laws see R. Sam Garrett, *Artificial Intelligence (AI) and Campaign Finance Policy: Recent Developments*, Congressional Research Service, updated 25 September 2024.

⁶⁶ NCSL, *Deepfakes in elections and campaigns*.

⁶⁷ Yubin Kim, 'The legal gray zone of deepfake political speech', *Cornell Journal of Law. & Public Policy: The Issue Spotter*, 24 October 2025.

⁶⁸ *Election Financing Amendment and Elections Amendment Act*, which added a new s 182.5 to *The Elections Act*. See also S Lambert, 'Ban on deepfakes, misinformation during elections among bills passed as Manitoba legislative session ends', *CBC News*, 6 November 2025. Accessed at <https://www.cbc.ca/news/canada/manitoba/elxn-disinformation-manitoba-bill-pass-9.6969688>

⁶⁹ Government of Canada, *Government of Canada Strengthens Election Protections Following Royal Assent of Bill C-25, The Strong and Free Elections Act*, Media release, 18 June 2026. Accessed at: <https://www.canada.ca/en/democratic-institutions/news/2026/06/government-of-canada-strengthens-election-protections-following-royal-assent-of-bill-c-25-the-strong-and-free-elections-act.html>. See new section 480.1, *Canada Elections Act*.

⁷⁰ See *Public Official Election Act* (South Korea) Article 82-8, which was inserted by amendment in 2023. See also 'Deepfakes spread across all crimes from sex offenses, fraud to election law violations', *Korea Times*, 18 September 2024. Accessed at: <https://www.koreatimes.co.kr/southkorea/law-crime/20240918/deepfakes-spread-across-all-crimes-from-sex-offenses-fraud-to-election-law-violations>

⁷¹ See *Parliamentary Elections Act 1954* (Singapore), s 61MA, which was inserted by the *Elections (Integrity of Online Advertising) (Amendment) Act 2024*.

⁷² *Parliamentary Elections Act 1954*, s 61N.

⁷³ *Regulation (EU) 2022/2065 on a Single Market for Digital Services*, Art 35. Very large online platforms and online search engines are defined in Art 33.

resembles existing persons is distinguishable through prominent markings.⁷⁴ In 2024 the European Commission published guidelines under the Act on the mitigation of systemic risks for electoral processes.⁷⁵ The EU's Artificial Intelligence Act (2024) also puts obligations on developers and deployers of AI systems that generate deepfakes.⁷⁶

In 2025 the Danish Government proposed amendments to the *Copyright Act* to address deepfakes in general.⁷⁷ The changes are designed to protect the public against the sharing of realistic, digitally generated imitations of a person's personal traits (such as looks and voice) without consent. Violations will not be criminalised. Focus will instead be on ensuring that people have the right to remove online content, and Danish and EU authorities can issue fines if the content is not removed. At the time of writing, it appears that the new laws have not been passed.

5. Other responses to political deepfakes

5.1 Deepfake counter measures

There are emerging deepfake counter measures which could help mitigate the potential adverse effects of political deepfakes. An overview of these counter measures is provided in Table 2.

Table 2: Brief overview of deepfake counter measures

Type	Features
Invisible watermarks	Watermarking is 'the process of concealing or embedding data behind an image or video that is invisible to the naked human eye'. ⁷⁸ They could be used to identify that particular content was generated by AI. ⁷⁹ This technology is still developing and not full-proof and requires industry co-operation.
Deepfake detection	Deepfake detection technologies are designed to reveal the origins of digital content, regardless of whether the content has been watermarked. ⁸⁰ This technology is still developing and not full-proof.
Prompt takedown	The adverse effects of political deepfakes may be reduced by their prompt removal from social media platforms, as occurred with the Irish presidential election.
Media fact-checking	The role of the media as a 'pillar of democracy', and its values of objectivity, accuracy and transparency, have become more important since the emergence of political deepfakes. ⁸¹
Public education	A 2024 senate committee report recommended that the government examine mechanisms to improve AI literacy, particularly in an electoral context. ⁸² However, it has been suggested that 'technological advances are becoming so sophisticated it is

⁷⁴ Art 35(1)(k).

⁷⁵ C/2024/3014 [Commission guidelines for providers of very large online platforms and very large online search engines on the mitigation of systemic risks for electoral processes pursuant to Article 35\(3\) of Regulation \(EU\) 2022/2065](#).

⁷⁶ [Regulation 2024/1689 laying down harmonised rules on artificial intelligence](#), Art 50(2).

⁷⁷ S Karttunen, *The Danish approach to copyright and deepfakes: A model for the EU?*, European Parliamentary Research Service, January 2026.

⁷⁸ P Kharvi, *Understanding the impact of AI-generated deepfakes*, pp. 2-9.

⁷⁹ B Chesney and D Citron, *Deepfakes: A looming challenge for privacy*, p. 1,814.

⁸⁰ United Kingdom Government, 'Deepfake detection technology', 26 March 2026. Accessed at <https://www.gov.uk/government/publications/deepfake-detection-technology/deepfake-detection-technology>.

⁸¹ P Raemy et al, [Deepfakes and journalism: Normative considerations and implications](#), *Journalism studies*, 26(14), 2025, pp. 1742-1744.

⁸² Australian Parliament, *Select Committee on Adopting Artificial Intelligence – Interim Report*, p. 35 (Rec 5).

Type	Features
	unreasonable to expect consumers and citizens to be able to “spot” deceptive imagery and voices’. ⁸³

5.2 Actions taken by digital platforms

At a conference in Munich in February 2024 27 technology companies including Google, Meta, and Microsoft adopted *A Tech Accord to Combat Deceptive Use of AI in 2024 Elections*.⁸⁴ The accord is a voluntary framework of principles and actions to advance 7 goals including prevention, detection, swift responses to incidents, and public awareness. In February 2025 the Brennan Center for Justice at New York University reviewed progress over the year and commented that ‘while some [companies] shared evidence of their actions, inconsistent follow-through, vague reporting, and a lack of independent verification made it challenging to assess real progress.’⁸⁵

5.3 Actions taken by electoral authorities

In the lead up to the 2019 federal election, the Australian Electoral Commission (AEC) ran a digital literacy campaign called ‘Stop and Consider’. The campaign was expanded for the 2025 federal election to address new topics such as AI.⁸⁶ The NSW Electoral Commission also had a ‘Stop and Consider’ campaign for the 2023 state election.⁸⁷ Some international electoral authorities are also exploring using deepfake detection technology. In January 2026 media in the United Kingdom reported that election officials had a pilot project with the Home Office to detect deepfakes that target candidates in the 2026 Scottish and Welsh elections.⁸⁸

6. Conclusion

Political deepfakes pose a new challenge to democratic institutions. The recent electoral law reforms in NSW include a ban on political deepfakes during an election period, with certain exceptions. Several issues arise in relation to these laws, including whether they are consistent with the implied freedom of political communication and how effective they will be in deterring political deepfakes. In addition to laws, combatting political deepfakes is likely to require a range of actions, including technological measures and education. The 2027 NSW election will provide an opportunity to see whether there are any cases of political deepfakes, what their impact will be, how the new laws are applied, and whether additional measures are required to safeguard elections in NSW.

⁸³ Inter-Parliamentary Union, ‘Dangers of deepfakes for parliamentarians’, 19 February 2024. Accessed at <https://www.ipu.org/news/news-in-brief/2024-02/dangers-deepfakes-parliamentarians>.

⁸⁴ Munich Security Conference, *A tech accord to combat deceptive use of AI in 2024 elections*, February 2024.

⁸⁵ A Ahmed et al, *Tech companies pledged to protect elections from AI – here’s how they did*, Brennan Center for Justice, 13 February 2025. Note that the Australian Communications and Media Authority monitors the activities of digital platforms under the *Australian Code of Practice on Disinformation and Misinformation*, which is not legally binding.

⁸⁶ Australian Electoral Commission, ‘AI & elections’. Accessed at <https://www.aec.gov.au/Elections/electoral-advertising/ai-and-elections.htm>

⁸⁷ NSW Electoral Commission, *Commission campaign to tackle election disinformation*, Media release, 21 February 2023.

⁸⁸ S Carrell, ‘Software tackling deepfakes to be piloted for Scottish and Welsh elections’, *The Guardian*, 9 January 2026. Accessed at <https://www.theguardian.com/technology/2026/jan/08/pilot-software-tackle-deepfakes-scottish-welsh-elections>